EFFECTS OF DISTRIBUTED AND MASSED PRACTICES OF VOCABULARY ASPECTS EMBEDDED IN A RESPONSE CARD ACTIVITY ON ACQUISITION, GENERALIZATION AND MAINTENANCE OF VOCABULARY KNOWLEDGE

DISSERTATION

Presented in Partial Fulfillment of the Requirement for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Madoka Itoi, M. A.

*****

The Ohio State University

2008

Dissertation Committee:
Professor Moira Konrad, Advisor
Professor Sheila Alber-Morgan
Professor Laurice Joseph

Approved by

Moira Konrad
Advisor
College of Education and Human Ecology
Extensive vocabulary knowledge is critical to a student’s academic and social performance in school. However, vocabulary is a complex skill that overlaps with many different skills and can be conceptualized with different dimensions. While there is a pressing need to provide effective and efficient vocabulary instruction at all levels, the perceived complexity of vocabulary creates confusion about how to operationalize vocabulary knowledge, how to consistently and systematically measure vocabulary improvements, and how to teach different dimensions of vocabulary. Available research suggests that students should (a) be actively engaged in instruction and (b) be allowed repeated practice opportunities. The current study examined the comparative effects of massed practice (MP) and distributed practice (DP) of aspects of vocabulary (i.e., spelling, writing definitions, identifying parts of speech, identifying synonyms, sentence writing) embedded in a response card activity on the vocabulary and spelling acquisition, maintenance, and generalization of five seventh grade students with and without disabilities enrolled in an inclusive language art classroom. The results were mixed, so definitive trends could not be determined. Despite the study’s limitations, there were several implications for future research and suggestions for classroom practice identified.
Dedicated to my family in Japan
ACKNOWLEDGMENTS

I wish to thank my advisor, Moira Konrad, for her intellectual and emotional supports in a consistent and warm manner that have made this dissertation possible. She has inspired me, encouraged me, and took the pressure off of me for these three years. Thank you for being the greatest role model for me.

I would like to thank Kim Cullen and her wonderful 7th grade students for helping me complete this project. They were the most flexible, nice, and fun people to work with.

I am grateful to my wonderful colleague, Ken Luu, for adjusting his busy schedule to help out the data collection.

Lastly, I would like to express my gratitude to my friends and students at the Japanese school, for always reminding me of why I like to teach and what it means to be an behavior analyst.
VITA

September 5, 1979 …………………… Born - Tokyo, Japan

March, 2002 ……………………….. B.A. in Education and Teaching English as Second Language

August, 2004 ……………………….. M.A. in Special Education and Applied Behavior Analysis

October, 2004 – August, 2005 ……… Behavior interventionist – Autism Center at Children’s Hospital, Columbus

September, 2005 – present ………… Graduate administrative and teaching assistant, The Ohio State University

PUBLICATIONS


FIELD OF STUDY

Major Field: Education

Area of Emphasis: Special Education and Applied Behavior Analysis
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapters:</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Instruction for Students with Disabilities</td>
<td>1</td>
</tr>
<tr>
<td>Research-to-Practice Gap</td>
<td>4</td>
</tr>
<tr>
<td>Vocabulary Instruction and Report from the National Reading Panel</td>
<td>10</td>
</tr>
<tr>
<td>Potential Contributions</td>
<td>12</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>12</td>
</tr>
<tr>
<td>Definitions</td>
<td>14</td>
</tr>
<tr>
<td>Delimitations</td>
<td>15</td>
</tr>
<tr>
<td>Review of the Literature</td>
<td>16</td>
</tr>
<tr>
<td>Active Student Responses (ASR)</td>
<td>16</td>
</tr>
<tr>
<td>Learning and ASR</td>
<td>16</td>
</tr>
<tr>
<td>ASR Strategies</td>
<td>17</td>
</tr>
<tr>
<td>Response Cards Studies</td>
<td>19</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>22</td>
</tr>
<tr>
<td>Development of Vocabulary and Relation to Other Skills</td>
<td>22</td>
</tr>
<tr>
<td>Vocabulary Instruction</td>
<td>24</td>
</tr>
<tr>
<td>Dimensions of Vocabulary Knowledge</td>
<td>29</td>
</tr>
<tr>
<td>Curriculum-based Measurement (CBM)</td>
<td>32</td>
</tr>
<tr>
<td>Massed Versus Distributed Practice</td>
<td>35</td>
</tr>
<tr>
<td>Practice Formats</td>
<td>35</td>
</tr>
<tr>
<td>Summary</td>
<td>40</td>
</tr>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iv</td>
</tr>
<tr>
<td>Vita</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xii</td>
</tr>
</tbody>
</table>
Method .................................................................................................................................................42
Setting ............................................................................................................................................42
Participants ....................................................................................................................................43
  Students ........................................................................................................................................43
  Experimenter .............................................................................................................................45
  Observer .......................................................................................................................................46
Independent variables .......................................................................................................................46
  Selection of Words and Materials .............................................................................................47
    Selection of Vocabulary Words .............................................................................................47
    Instructional Materials ...........................................................................................................49
Dependent measures .........................................................................................................................52
  Vocabulary Acquisition Measures ......................................................................................53
  Spelling Acquisition Measures ...........................................................................................54
  Generalization Measures ......................................................................................................56
  Maintenance Measures ...........................................................................................................56
  Social Validity ..........................................................................................................................57
Experimental Design .......................................................................................................................57
Data Collection Procedures ...........................................................................................................58
  Probe Procedure ..................................................................................................................58
  Instructional Procedure .......................................................................................................60
Procedures for Obtaining Reliability of Data ...............................................................................64
  Procedure for Obtaining Procedural Integrity ......................................................................65
Procedure for Obtaining Teacher’s and Students’ Opinions .....................................................65
Results ................................................................................................................................................68
Believability of Data .......................................................................................................................69
  Procedural Integrity .................................................................................................................69
  Interobserver Agreement (IOA) .............................................................................................69
What are the Effects of Distributed Versus Massed Practices Embedded in a Response-Card
Activity on the Vocabulary Performances as Measured by Accuracy of Definition Writing? ....
.......................................................................................................................................................70
  Mean accuracy of definition writing ...................................................................................70
  Student 1 (Jessie) ..................................................................................................................72
  Student 2 (David) ..................................................................................................................74
  Student 3 (Rachel) ...............................................................................................................77
  Student 4 (Kelly) ....................................................................................................................80
  Student 5 (Sam) ......................................................................................................................83
What are the Effects of Distributed Versus Massed Practices Embedded in a Response-Card
Activity on the Vocabulary Performances as Measured by Identification of Parts of Speech?...87

Student 1 (Jessie) ..................................................................................................................87
Student 2 (David) ..................................................................................................................90
Student 3 (Rachel) ...............................................................................................................93
Student 4 (Kelly) ................................................................................................................97
Student 5 (Sam) ..................................................................................................................100

What are the Effects of Distributed Versus Massed Practices Embedded in a Response-Card Activity on the Vocabulary Performances as Measured by Accuracy of Identification of Synonym? .................................................................................................................................103

Mean accuracy of identification of synonyms ............................................................................103

Student 1 (Jessie) ................................................................................................................105
Student 2 (David) ................................................................................................................108
Student 3 (Rachel) ..............................................................................................................111
Student 4 (Kelly) ................................................................................................................114
Student 5 (Sam) ..................................................................................................................117

What are the Effects of Distributed Versus Massed Practices Embedded in a Response-Card Activity on the Spelling Performances as Measured by the Number of Correctly Spelled Words? ........................................................................................................................................120

Mean accuracy of spelling ........................................................................................................120

Student 1 (Jessie) ................................................................................................................120
Student 2 (David) ................................................................................................................125
Student 3 (Rachel) ..............................................................................................................128
Student 4 (Kelly) ................................................................................................................131
Student 5 (Sam) ..................................................................................................................134

What are the Effects of Distributed versus Massed Practices Embedded in a Response-Card Activity on the Spelling Performances as Measured by the Percent of Correct Letter Sequences (CLS)? .............................................................................................................................137

Student 1 (Jessie) ................................................................................................................137
Student 2 (David) ................................................................................................................139
Student 3 (Rachel) ..............................................................................................................141
Student 4 (Kelly) ................................................................................................................141
Student 5 (Sam) ..................................................................................................................145

What are the Effects of Distributed Versus Massed Practices Embedded in a Response-Card Activity on the Generalization of Vocabulary Knowledge as Measured by the Ratings of Novel Sentences Depicted in a Four-Point Likert Scale? ..............................................................145

Mean ratings for generalization ................................................................................................145

Student 1 (Jessie) ................................................................................................................146
Student 2 (David) ................................................................................................................146

ix
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 List of Vocabulary Words for Each Condition</td>
<td>48</td>
</tr>
<tr>
<td>4.1 Students’ Opinions about MP and DP of Vocabulary</td>
<td>166</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Instructional slides</td>
<td>50</td>
</tr>
<tr>
<td>3.2 Sequence of instructional prompt during the response card practice</td>
<td>63</td>
</tr>
<tr>
<td>4.1 Mean accuracy of definition writing</td>
<td>71</td>
</tr>
<tr>
<td>4.2 Accuracy of definition writing for student 1 (Jessie)</td>
<td>73</td>
</tr>
<tr>
<td>4.3 Cumulative accuracy of definition writing for student 1 (Jessie)</td>
<td>74</td>
</tr>
<tr>
<td>4.4 Accuracy of definition writing for student 2 (David)</td>
<td>75</td>
</tr>
<tr>
<td>4.5 Cumulative accuracy of definition writing for student 2 (David)</td>
<td>76</td>
</tr>
<tr>
<td>4.6 Accuracy of definition writing for student 3 (Rachel)</td>
<td>78</td>
</tr>
<tr>
<td>4.7 Cumulative accuracy of definition writing for student 3 (Rachel)</td>
<td>79</td>
</tr>
<tr>
<td>4.8 Accuracy of definition writing for student 4 (Kelly)</td>
<td>81</td>
</tr>
<tr>
<td>4.9 Cumulative accuracy of definition writing for student 4 (Kelly)</td>
<td>82</td>
</tr>
<tr>
<td>4.10 Accuracy of definition writing for student 5 (Sam)</td>
<td>84</td>
</tr>
<tr>
<td>4.11 Cumulative accuracy of definition writing for student 5 (Sam)</td>
<td>85</td>
</tr>
<tr>
<td>4.12 Accuracy of identification of parts of speech for student 1 (Jessie)</td>
<td>88</td>
</tr>
<tr>
<td>4.13 Cumulative accuracy of identification of parts of speech for student 1 (Jessie)</td>
<td>89</td>
</tr>
<tr>
<td>4.14 Accuracy of identification of parts of speech for student 2 (David)</td>
<td>91</td>
</tr>
</tbody>
</table>
4.15. Cumulative accuracy of identification of parts of speech for student 2 (David) .......92
4.16. Accuracy of identification of parts of speech for student 3 (Rachel) .................95
4.17. Cumulative accuracy of identification of parts of speech for student 3 (Rachel) ....96
4.18. Accuracy of identification of parts of speech for student 4 (Kelly) ....................98
4.19. Cumulative accuracy of identification of parts of speech for student 4 (Kelly) ....99
4.20. Accuracy of identification of parts of speech for student 5 (Sam) .......................101
4.21. Cumulative accuracy of identification of parts of speech for student 5 (Sam) ....102
4.22. Mean accuracy of identification of synonyms ..................................................104
4.23. Accuracy of identification of synonyms for student 1 (Jessie) .........................106
4.24. Cumulative accuracy of identification of synonyms for student 1 (Jessie) .........107
4.25. Accuracy of identification of synonyms for student 2 (David) .........................109
4.26. Cumulative accuracy of identification of synonyms for student 2 (David) .........110
4.27. Accuracy of identification of synonyms for student 3 (Rachel) .......................112
4.28. Cumulative accuracy of identification of synonyms for student 3 (Rachel) .......113
4.29. Accuracy of identification of synonyms for student 4 (Kelly) .........................115
4.30. Cumulative accuracy of identification of synonyms for student 4 (Kelly) .........116
4.31. Accuracy of identification of synonyms for student 5 (Sam) .........................118
4.32. Cumulative accuracy of identification of synonyms for student 5 (Sam) .........119
4.33. Mean accuracy of spelling ..............................................................................121

xiv
4.34. Accuracy of spelling for student 1 (Jessie) ..............................................................................123
4.35. Cumulative accuracy of spelling for student 1 (Jessie) ...........................................................124
4.36. Accuracy of spelling for student 2 (David) ..............................................................................126
4.37. Cumulative accuracy of spelling for student 2 (David) ..........................................................127
4.38. Accuracy of spelling for student 3 (Rachel) ............................................................................129
4.39. Cumulative accuracy of spelling for student 3 (Rachel) .........................................................130
4.40. Accuracy of spelling for student 4 (Kelly) ...............................................................................132
4.41. Cumulative accuracy of spelling for student 4 (Kelly) .............................................................133
4.42. Accuracy of spelling for student 5 (Sam) ................................................................................135
4.43. Cumulative accuracy of spelling for student 5 (Sam) .............................................................136
4.44. Percent correct letter sequences (CLS) for student 1 (Jessie) .................................................138
4.45. Percent correct letter sequences (CLS) for student 2 (David) ................................................140
4.46. Percent correct letter sequences (CLS) for student 3 (Rachel) ...............................................142
4.47. Percent correct letter sequences (CLS) for student 4 (Kelly) ..................................................143
4.48. Percent correct letter sequences (CLS) for student 5 (Sam) ...................................................144
4.49. Mean accuracy rating of novel sentences (Generalization) ..................................................147
4.50. Accuracy rating of novel sentences (Generalization) for student 1 (Jessie) ...........................149
4.51. Cumulative accuracy rating of novel sentences (Generalization) for student 1 (Jessie) ......150
4.52. Accuracy rating of novel sentences (Generalization) for student 2 (David) .........................152
4.53. Cumulative accuracy rating of novel sentences (Generalization) for student 2 (David) ....153
4.54. Accuracy rating of novel sentences (Generalization) for student 3 (Rachel)..................155
4.55. Cumulative accuracy rating of novel sentences (Generalization) for student 3 (Rachel).....156
4.56. Accuracy rating of novel sentences (Generalization) for student 4 (Kelly)..................158
4.57. Cumulative accuracy rating of novel sentences (Generalization) for student 4 (Kelly) ......159
4.58. Accuracy rating of novel sentences (Generalization) for student 5 (Sam).....................161
4.59. Cumulative accuracy rating of novel sentences (Generalization) for student 5 (Sam).......162
CHAPTER 1

INTRODUCTION

Instruction for Students with Disabilities

The Individuals with Disabilities Education Improvement Act, or IDEA, specifies that students with disabilities are entitled to a free appropriate public education accomplished through an individualized education program (IEP) and that students with disabilities should be educated with students without disabilities to the maximum extent possible (Heward, 2006). These two principles of IDEA (i.e., free appropriate education and the least restrictive environment) resulted from the efforts of individuals committed to expanding the educational rights of students with disabilities and the practice of inclusive education. No Child Left Behind (NCLB) advanced this movement a step further by mandating that all students, including those with disabilities, participate in state and district assessment and that evidence-based, effective instruction be provided to all students including those with disabilities (NCLB, 2001).

Reflecting these laws, provision of educational services for students with disabilities occurs in various environments today. According to the National Center for Education Statistics, 49.9% of the students from ages 6 to 21 who were served under IDEA spent 80% of their school day in the general education classroom in the 2003-2004
school year, followed by resource room or special classrooms within the public school. In addition, the proportion of students who spend the majority of their day in general education classrooms has increased by 5% in a decade and the same trend is also expected to continue (U. S. Department of Education, 2005). Therefore, under the current legal practice, one of the major challenges of special education is to ensure effective instruction for students with disabilities across placement options.

Anderson-Inman, Paine, and Deutchman (1984) suggested that there were three types of skills required for students with developmental delay and other low performing students to succeed in school. The skills are (a) social skills sufficient to allow appropriate interactions with teachers and peers, (b) basic academic skills sufficient to keep up with the curriculum, and (c) academic support skills needed to benefit from classroom instruction and subsequently, to demonstrate that learning has occurred. Despite the fact that more and more students with disabilities are now at least partially included in general education settings, there are certain characteristics of these students that call for additional, intensive, and specialized instruction (i.e., special education) in the aforementioned three areas (Heward, 2006). Specifically, students with disabilities may exhibit lower or shorter levels of attention to task at hand (Mooney, Ryan, Uhing, Reid, & Epstein, 2005; Lloyd, Hallahan, Kosiewicz, & Kneedler, 1982) and may require more extensive or frequent prompts to stay on task; they may lack some prerequisite or basic skills to engage in tasks, requiring more intensive individualized instruction (Callicott & Park, 2003; Martin & Manno, 1995; Montague, 2007); and they may require
different procedures such as visual prompts to increase the likelihood of responsiveness to instruction (Dooley, Wilczenski & Torem, 2001) partly due to lack of motivation (Montague).

Characteristics of the general education environment may present additional challenges that interfere with the effort to provide appropriate and necessary services to students with disabilities in inclusive settings (Anderson-Inman et al., 1984; Hamilton, Seibert, Gardner, & Talbert-Johnson, 2000). For example, with more students physically present in the classroom, teachers are more likely to employ lecture or other group-oriented instructional methods. The instructional formats do not readily lead to intensive one-to-one support. In terms of pre-service training, general education teachers are less likely to be trained in behavior management or instructional strategies for students with disabilities (Lewis et al., 1999). As a result, general education teachers may not always be able to provide the level of structure and clarity in the delivery of instruction as special education teachers usually do (Anderson-Inman et al., 1984), or provision of intensive support in general education settings may sacrifice loss of actual teaching time and integrity of interventions (Neef, 1995).

The perceived mismatch between the needs of students with disabilities and the characteristics of general education classrooms may result in negative academic performances of students with disabilities. Research suggests that children who fail to learn to read by the first grade tend to fall further behind in reading and as well as in other general academic areas. For example, Fletcher and colleagues (1994) conducted a longitudinal study of 199 children, ages 7 to 9 years old. The results suggested that about 74% of children diagnosed as having learning disabilities (LD) continued to have
problems in the ninth grade. A similar trend is evident among students with emotional and behavioral disorders (E/BD), who typically show chronically below-grade-level performance compared to peers (cf., Cullinan, Evans, Epstein, & Ryser, 2003). Also, studies show that 20 to 70% of students with ADHD have other disabilities that adversely affect their academic performances (Mayes, Calhoun, & Crowell, 2000). Perceived academic difficulties during school years result in negative post-school outcomes. Nearly 20% of youth with learning disabilities (LD) are not engaged in work or education after leaving high school. About 15% of students who are employed earn less than federal minimum wage. In addition, youth with LD and other health impairments are more likely to be involved with criminal justice system (Wagner, Newman, Cameto, & Levine, 2005). From these statistics, the importance of academic instruction increases so that the reading difficulties (and hence, other academic deficits) of students with disabilities will be minimized as early as possible.

*Research-to-Practice Gap*

In the field of special education, there are ample examples of behaviorally-oriented effective instructional practices validated through repeated empirical demonstrations. Cooper (1982), with everyday classroom examples, discussed how ABA (Applied Behavior Analysis) technologies can benefit students and, ultimately, empower generations of teachers by providing basis for guiding daily instruction. He raised three contributions ABA can offer to education: (a) a research foundation for the systematic replication of environmental manipulation to change academic and social behaviors, (b) an emphasis on direct and repeated measurement of target behaviors in the classroom, and (c) the development of single-subject research designs that enable
classroom teachers to become practitioners of applied research. That is, teachers who employ the three approaches of ABA do not have to rely solely on random trial and error learning, verbal reports, and informal observations to discover principles of learning, and they can effectively evaluate their instruction by looking at their data and adjust their instruction when necessary (cf. Heward, 2005). There are more specific examples of “technologies” that have been established and recognized as effective practices in education. For example, Becker and Engelmann (1996) provided an excellent specific example of how carefully designing and packaging good curriculum, classroom procedures, and training procedures based on behavioral principles could remediate performances of many academically disadvantaged students in major achievement measures such as word knowledge, spelling, language, math computation, reading comprehension, and problem solving, as well as affective measures of cooperation, self-esteem, or intellectual achievement. Another example is the use of functional analysis and functional assessments. Functional assessment (FA) of the problem behaviors of students with disabilities is now recognized as a component of best practices recommendations (Reid, 2000) and a necessary step of positive behavior support (Heward, 2005). Furthermore, after the passage of the Individuals with Disabilities Education Act (IDEA) in 1997, functional behavior assessment is a required procedure for all students with disabilities in public educational programs under specific conditions (Scott, Liaupsin, Nelson, & McIntyre, 2005). Also, schoolwide positive behavior support (SWPBS) has been implemented at school- and district-level to address and prevent disciplinary issues and to teach appropriate alternative behaviors and demonstrating its effectiveness (Anderson & Kincade, 2005; Horner, Sugai, Lewis-Palmer, & Todd, 2001).
However, as frequently as people have documented potential utility of behavioral principles in improving education, there have been frustrated, perplexed, or sometimes devastating accounts about why ABA is failing to positively influence (or failing to make as big an impact as it should on) the behaviors of people involved in education, despite potential benefit. Carnine (2000), for example, reported how empirical evidence documenting the effectiveness of the Direct Instruction (DI) approach was ignored at the national level. Lindsley (1992) rather pessimistically attributed the unsuccessful adoption of behavioral instructional strategies to current academic and/or political contingencies such as availability of money, social interest, popular myths, or politics. Baer (1992) attributed the same phenomenon to insufficient attention paid by behavioral researchers to contingencies in “natural” educational settings: the effects of various ecological or extraneous variables, curriculum development, and philosophy adopted by powerful political/social groups (cf. Axelrod, 1992). These phenomena continue to exist despite ample demonstrations of effective instructional and behavior management strategies from special education and applied behavior analysis in many different classroom arrangements, including general education classrooms. This incongruence between what researches have already suggested about effective practice and what is actually practiced in classrooms is known as the research-to-practice gap (Carnine, 1997; Heward, 2003).

When existing “effective” evidence-based practices are not used in actual classrooms, they face social invalidity or rejection from the consumers due to potential problems in at least one of the following areas: their goals, means or procedures they employ, or the produced effects (Wolf, 1978). Schwartz and Baer (1991) argued for the
need to acknowledge various levels of consumers and to ask the right questions or identify the right behavioral indicators of social validity for each level. With the research-to-practice gap, the research evidence may not be answering the “right” questions that the direct consumers of these practices (i.e., teacher, instructional aides, caregivers, or therapists) are most interested in. In order to refine the research questions and findings to address the research-to-practice gap, Carnine (1997) suggested a different set of important considerations. He argued that consumers evaluate results of educational research in terms of trustworthiness (the confidence with which a given set of findings can be acted upon by practitioners, policy makers, publishers, and other knowledge consumers), usability (the likelihood that research will be used by those who have the responsibility for making decisions), and accessibility (the ease and quickness with which practitioners can obtain necessary information related to certain goal).

The existing research-to-practice gap highlights the unique nature of educational instructional technologies. They are evaluated using multiple criteria. First, they have to produce outcomes or some evidence that each student learned and used new knowledge and skills (Heward, 2003). In addition to this basic criterion, an outcome may become more meaningful if the student can apply the knowledge or the skill to under untrained set of stimuli (i.e., generalizability) or for extended period of time after the initial training is terminated (i.e., maintenance) (Peterson, Young, Salzberg, West, & Hill, 2006). Third, they need to pass the social validity test from various consumers, by providing trustworthy, usable, and accessible knowledge base (Carnine, 1997).

There are two types of experiments that can enhance the levels of trustworthiness, usability, and accessibility of educational technologies: component
analysis and parametric analysis. Component analyses are used when the intervention is “packaged” or have one or more components and the researcher is interested in discovering which component is most responsible for the perceived changes in the dependent variable(s) (Cooper, Heron, & Heward, 2007). With the information obtained from component analyses, teachers can save the effort of implementing all components of an intervention and focus only on the most influential part, therefore enhancing the levels of trustworthiness, usability, and accessibility of educational technologies. Parametric analyses discover differential effects of a range of values of a single independent variable over the dependent variable(s) (Cooper et al., 2007). By conducting this type of analysis, practitioners gain the basis for predicting the intensity of effects on the dependent variable(s) associated with certain level with which they implement the independent variable. For example, Prator, Hogan, and Miller (1992) conducted a study that involved both component and parametric analyses. They used a multiple baseline across settings design to examine the effects of a self-monitoring package involving self-recording, audio tones, and a poster. Also, researchers varied the intervals with which the student heard tones that signaled when to self-record his on-task behavior. The results suggested that although the highest level of on-task behavior was observed under the condition with all intervention components, frequency of tones did not greatly affect the level of on-task behavior. Also, the level of academic performance did not differ with varying frequency of tones nor intervention components. Courson (1989) examined the effects of long or short-form of guided notes on the next-day quiz scores of 19 seventh grade students with learning disabilities or academically at-risk students. The long- or short-forms of guided notes did not produce significant difference of the students’ quiz scores. Therefore, this
study is considered as a parametric analysis that informed the practitioners that providing guided notes alone could benefit students, regardless of the amount of writing they do during the lecture.

More studies are needed that inform practitioners about trustworthiness, usability, and accessibility of available “evidence-based” instructional strategies. Among several “evidence-based” instructional strategies, drill and practice strategies and active student responses (ASR) have high potential for wide classroom application. However, some multi-component instructional methods do not yield to omission of a component, making component analyses unrealistic. One example is the guided note strategy that involves notes, lectures, and overhead or PowerPoint presentations. Omitting some of the components may not necessarily result in good performance. For example, White (1991) compared students’ own notes and guided notes with and without overhead presentation on accuracy of notes taken by eight high-school students with learning disabilities. Although guided notes without overhead presentation improved note accuracy of these students compared to their own notes, overhead presentation produced greater improvement in the accuracy of notes. Unlike component analyses, parametric analyses answer more practical questions often posed by practitioners, such as “How often the intervention should be implemented?” or “What is the least possible frequency with which this intervention can be implemented, while still producing academic improvement in my students?” Taken together, a parametric analysis of a vocabulary intervention using drills and an ASR strategy may have a high application potential.
Students with disabilities experience difficulty in many skill areas; however, literacy skills may be most critical because they are necessary for learning in other areas. Researchers have identified that approximately 50% of students with learning disabilities are considered “non-responders” for generally effective interventions; as a result, these students continue to have reading difficulties (Al Otaiba & Fuchs, 2006). Al Otaiba and Fuchs conducted a longitudinal study to identify characteristics of students who are non-responsive to early literacy intervention conducted by their kindergarten and first grade classroom teachers on 104 students. Students were grouped into four groups based on the timing of their participation to the best practice instruction: (a) in kindergarten and first grade, (b) in kindergarten only, (c) in first grade only, and (d) in neither year. After two years of instruction, multivariate analysis of variance and discriminant function analysis were used to identify the students as always responsive, sometimes (i.e., one of the two years) responsive, or non-responsive. Results indicated that measures of problem behavior, verbal memory, sentence imitation, syntactic awareness, vocabulary, naming speed, and segmentation reliably differentiated responsive and non-responsive students. Measures of naming speed, vocabulary, sentence imitation, problem behavior, and amount of intervention were also identified as predictors of non-responsive and responsive students. In addition, among five students who were identified as non-responsive to intervention during kindergarten, first grade, or both kindergarten and first grades, four students had IEPs in reading in the third grade. Results of this study and other studies (e.g., Torgesen & Davis, 1996) indicate the importance of strong vocabulary skills as one of the predictors of reading achievement. In other words, an achievement
gap that exists during kindergarten or first grade widens as students proceed through their educational careers, partly because of differences in vocabulary skills.

The report from *The National Reading Panel* identified five areas that correlate with positive student reading outcomes: phonemic awareness, phonics, reading fluency, comprehension, and vocabulary instruction (NRP, 2000). Although it may be less related to the acquisition of beginning code-based reading skills, vocabulary knowledge plays an increasingly significant role in text comprehension as students progress through the elementary grades and as they are exposed to more complex texts (Coyne, McCoach, & Kapp, 2007). Therefore, in order to reduce and prevent the achievement gap of students with disabilities, teachers should provide appropriate vocabulary instruction so that those students can get onto the right trajectory of vocabulary and reading development and maintain that growth (Apthorp, 2006).

Despite the alleged importance of vocabulary knowledge on students’ achievement in literacy and other areas, available suggestions about vocabulary instruction remain unspecific. For example, the NRP report noted several implications of vocabulary instruction in relation to teaching reading. First, vocabulary should be taught both directly and indirectly. Second, repetition and multiple exposures to vocabulary items should be provided. Third, contextual learning, incidental learning, and computer-assisted learning are found to enhance acquisition of vocabulary. Direct instruction of vocabulary should restructure the tasks if necessary and embed active student engagement. Finally, dependence on a single method of vocabulary instruction should be avoided. Although these five suggestions imply critical considerations for practitioners when designing activities for vocabulary instruction, they do not necessarily
mention the specific activities. Reflecting this point, the report highlighted the lack of empirical research on certain methods of vocabulary instruction or combination of methods of instruction and their relation to reading achievement (NRP, 2000; cf. Apthorp, 2006). In summary, not only did it emphasize the importance of vocabulary, the NRP’s report also echoed the frequently-voiced concern from practitioners and researchers in saying that classrooms await research on optimal conditions or combination of methods of vocabulary instruction that meet the five suggestions from its original report.

**Potential Contributions**

Vocabulary instruction is often done incidentally especially for young students through exposure to different stories or activities (e.g., Braham & Lynch-Brown, 2002; Coyne, McCoach, & Kapp, 2007); however, for older students, there has been limited empirical investigation regarding how teachers should teach vocabulary skills. This study is an attempt to identify instructional conditions to teach complex vocabulary content while maintaining the efficacy of the instructional arrangements. If a short instructional session using the available “best practices” can teach different vocabulary content to seventh grade students with and without disabilities, the results can answer many questions about how we should teach vocabulary to older students.

**Purpose of the Study**

The purpose of this study was to examine the effects of distributed and massed practice embedded in an ASR-rich instructional context on the acquisition, maintenance, and generalization of language arts vocabulary words of seventh grade students with disabilities (mild mental retardation, speech and language disability, and ADHD) and without disabilities. More specifically, this study sought to answer the following research
questions:

1. What are the effects of distributed versus massed practices embedded in a response-card activity on the vocabulary performances of seventh grade students with and without disabilities in an inclusive language arts classroom as measured by
   (a) the accuracy of definition writing,
   (b) the accuracy of identification of parts of speech, and
   (c) the accuracy of recognition of synonyms illustrating the vocabulary words?

2. What are the effects of distributed versus massed practices embedded in a response-card activity on the spelling performances of seventh grade students with and without disabilities in an inclusive language arts classroom as measured by
   (a) the number of correctly spelled words and
   (b) the percent of correct-letter-sequence (CLS)?

3. What are the effects of distributed versus massed practices embedded in a response-card activity on generalization of words in the form of novel sentence writing using the vocabulary words of seventh grade students with and without disabilities in an inclusive language arts classroom as measured by the accuracy of vocabulary words usage in novel sentences depicted in a four-point likert scale?

4. What are the effects of distributed versus massed practices embedded in a response-card activity on maintenance of words of seventh grade students
with and without disabilities in an inclusive language arts classroom as measured by the accuracy of definition writing?

5. How socially valid are the distributed and massed practices embedded in a response-card activity as measured by

(a) the comparison of above dependent measures between the five students with disabilities and students without disabilities,

(b) the ratings of distributed or massed practice formats by students with and without disabilities regarding their vocabulary performances, and

(c) the ratings of distributed or massed practice formats by the classroom teacher of seventh grade students with and without disabilities regarding the usability of the intervention procedures?

Definitions

Although the definitions of the following terms differ in many educational studies, this study adapted the definitions described in this section. Familiarity to these definitions may enhance understanding of this study.

Aspects of vocabulary: Stands for spelling, definition, parts of speech, and synonyms of a word.

Dimensions of vocabulary: Five subcategories of vocabulary. They refer to phonology, morphology, syntax, semantics, and pragmatics.

Distributed practice: Practicing four aspects of vocabulary of a word before moving to the practice of the aspects of the next word. More specifically, the sequence of practice goes as follows: spelling (word 1), definition (word 1), synonym (word
Massed practice: Practicing one of the four aspects of vocabulary in a chunk before practicing another aspect. More specifically, the practice sequence goes as follows: spelling (words 1 through 5), definition (words 1 through 5), identification of synonym and parts of speech (words 1 through 5).

Delimitations

This study has some delimitations that may affect generalization of the results to individuals and settings studied in the current investigation. First, this study employed a single-subject design instead of statistical analysis with a large number of participants as an attempt to make an explicit judgment about the events under investigation (Baer, Wolf, & Risley, 1968). Therefore, the results may not be applicable to a wider population of students.

Second, vocabulary “content” in this study refers to a combination of distinct behaviors such as spelling, writing definition, identifying part of speech and synonyms, and writing novel sentences for a word. Since these “contents” differ in the form of response requirements, the results need to be analyzed for each aspect and therefore, drawing a general conclusion may be difficult.
CHAPTER 2

REVIEW OF LITERATURE

The purpose of this study was to examine the effects of distributed and massed practice embedded in an ASR-rich instructional context on the acquisition, maintenance, and generalization of language arts vocabulary words of seventh grade students with disabilities (mild mental retardation, speech and language disability, and ADHD) and without disabilities. In the following sections, I will provide the rationale behind the selection of the independent variables used in this study by summarizing the implications from previous studies about learning and vocabulary instruction. More specifically, I will summarize studies on active student response (ASR), ASR strategies, vocabulary instruction, curriculum-based measurement (CBM), and massed and distributed practice.

Active Student Responses (ASR)

Learning and ASR

Numerous research studies have repeatedly demonstrated that students’ achievement is enhanced when students actively engage in the instructional materials and activities (Berliner & Rosenshine, 1976; Heward, 1994; Sterling, Barbetta, Heward, & Heron, 1997). One way to increase active engagement is to increase the number of instructional episodes in which students participate. An instructional episode, also known as a learning trial, typically involves the three-tem contingency, which consists of an
instructional antecedent (e.g., teacher-posed question or question on a worksheet), a student response (e.g., verbal or written answer to a teacher-posed question), and an instructional consequence (e.g., praise or corrective feedback). Heward (1994) argues that increasing the number of learning trials provided for students increases the frequency of active student responses (ASR) and thus improves students’ achievement. According to Heward, three major benefits of increased frequency of ASR include (a) increased opportunities for students to practice target skills with receive feedback, (b) immediate and on-going feedback for teachers regarding the effectiveness of their instruction, and (c) increased on-task behaviors. Therefore, incorporating frequent ASR is the first step to ensure students’ achievement for any academic behavior.

ASR Strategies

Students’ active responses toward instructional antecedents can take many forms such as words read, problems answered, words or sentences written, worksheet questions answered, questions asked, or comments made (Heward, 1994; Shabani & Carr, 2004). Because ASR allows variety in its forms on the students’ side, teachers can design lessons and develop materials at their liberty based on classroom or student factors, or on the resources available. While there are many possible ways for teachers to incorporate opportunities for ASR during their daily instruction, Heward suggested three “low-tech” strategies for increasing the frequency of ASR in group instruction contexts. These strategies are termed “low-tech” for the availability of materials and applicability to many different activities typically observed in group instruction settings, and include response cards, choral responding, and guided notes (Heward).
Among the three low-tech ASR strategies, choral responding, in which all students orally respond in unison, and response cards, where each student holds up a card or board depicting their chosen answer, may have superior usability (Godfrey, Grisham-Brown, Schuster & Hemmeter, 2003; Heward, 1994). These two strategies (including two forms of response cards: pre-printed and write-on) are characterized by relatively simple skill requirements such as verbally responding, identifying, or writing words on a white board in response to teacher-posed questions, as well as potential for use in various subject areas and activities without extensive preparation on the teacher’s part.

Godfrey and her colleagues (2003) examined the comparative effects of pre-printed response cards, choral responding, and hand-raising on participation, on-task behavior, and inappropriate behavior of five preschool children who had difficulty attending to whole-group instruction during morning circle. Using an alternating treatment design (Cooper et al., 2007), the children responded to teacher-posed questions about the calendar using one of the three techniques for the duration of 25 sessions. The results indicated that response cards produced the highest number of active responses, the highest percentage of intervals of on-task behavior, and the lowest number of inappropriate behaviors across all participants. Although the results do not provide any information regarding the effects of response cards on the accuracy of students’ responses, they may support the effectiveness of response cards compared to other ASR strategies. Authors suggested the superior effects of response cards may be due to the complexity of required modes for learning. While response cards require kinesthetic, auditory, and
visual modes, hand-raising requires kinesthetic and auditory and choral responding requires only auditory (Godfrey et al., 2003).

It is important to consider modes of learning when a teacher designs curriculum for skill acquisition. For example, when the terminal behavior involves writing such as essay writing or completing math problems on a worksheet, response cards, especially write-on response cards, can provide ample opportunities to practice the terminal behavior compared to choral responding, which may provide students opportunities to engage in irrelevant behaviors (i.e., oral responding) repeatedly. Even if the terminal behavior does not match with the responses used with write-on response cards, write-on response cards can incorporate spelling instruction and involve responses other than simple recognition (Heward, 1994). Therefore, together with various learning modes involved, and practice opportunities of potentially relevant or complex behavior, response cards, particularly write-on response cards, may have the most practical usability among the three low-tech ASR strategies.

Response Cards Studies

To date, 14 empirical studies are available demonstrating the effects of response card strategy with many different subject areas with various characteristics. However, these studies are rather limited in terms of the context and skills with which response cards are used. Among the 14 existing empirical studies, there exists one study that used response cards for specific skill acquisition (i.e., telling time) (Horn, Schuster, & Collins, 2006). In this study, lessons using pre-printed response cards and hand-raising were examined to compare effects on mean percentage of active responding, on-task and inappropriate behaviors, and correct responding of identification of time by three middle
school students with mild to moderate disabilities. Response cards lessons resulted in more active responding and on-task behaviors, fewer inappropriate behaviors, and generally more correct responding on the task except for one student who responded equally well in both conditions.

Meanwhile, twelve studies used response cards during a review activity of lecture content and examined their effects on post-lecture quizzes (Kellum, Carr, & Dozier, 2001; Maheady, Michielli-Pendl, Mallette, & Harper, 2002; Marmolejo, Wider, & Bradley, 2004; Narayan, Heward, Gardner, Courson, & Omness, 1990), next-day quizzes (Cavanough, Heward, & Donelson, 1996; Gardner, Heward, & Grossi, 1994), weekly quizzes (Cavanough et al., 1996; Christle & Schuster, 2003; Clayton & Woodard, 2007; Gardner et al., 2004), and a post-test (Shabani & Carr, 2004).

Although these review activity studies clearly demonstrate the effectiveness of both preprinted and write-on response cards with students at different age levels and in the contexts of different subject matter, specific requirements for optimal use remain unknown. For example, instructional trials that utilized response cards are sometimes interspersed between existing activities (Armendariz, & Umbreit, 1999; Berrong, Schuster, Morse, & Collins, 2007; Gardner et al., 1994; Godfrey et al., 2003; Maheady et al., 2002; Marmolejo et al., 2004; Narayan et al., 1990; Shabani & Carr, 2004;) or maintenance items (Cavanough et al., 1996). In other cases, a review activity was developed especially to allow for response cards opportunities (Christle & Schuster, 2003; Clayton & Woodard, 2007; Davis & O'Neill, 2004; Kellun et al., 2001). Furthermore, in most studies, general knowledge (e.g., science, social studies, or psychology) was targeted in most of the review studies, and a few studies targeted a
specific skill (e.g., telling time) or class of specific skills (e.g., grammar and writing-related skills; Davis & O’Neill, 2004; calendar, Berrong et al., 2007; Godfrey, 2003). In all of these studies, the generalizability of these targeted knowledge or skills was not directly assessed using different forms used in the instructional context. That is, although the targeted knowledge about social studies or science or skills such as telling time can be assessed in many different forms, the assessment methods were limited to either worksheets or verbal responses and the assessment was done in a manner as to restrict students’ free responses.

From these studies, response cards, both preprinted and write-on, have shown to improve students’ quiz performances and study-related behaviors such as on-task or off-task behaviors when used in the context of review of lecture content designed in various format. In contrast, the number of demonstrations that used response cards in skill acquisition is limited to only one study that used preprinted response cards and the skill targeted for acquisition was highly specific. Given the higher usability of write-on response cards, additional demonstration of this strategy with skill acquisition is awaited. In addition, skills targeted for acquisition should be assessed in many different ways so that both response and stimulus generalizations can be assessed. In this respect, there is no existing study that focused on the use of write-on response cards on vocabulary acquisition, a skill that serves as a basis for other learning in many other areas and that can be assessed in many different formats. Also, optimal formats of vocabulary instruction using write-on response cards should be examined.

In summary, among available ASR strategies, write-on response cards have both empirical validation and practical usability. The response card strategy conforms to
variety of classroom settings and activities. However, previous demonstrations of effectiveness of response cards have been done with reviews of previously learned content and only one study has examined the effectiveness of using response cards for acquisition of concept or skills (Horn et al., 2006). Also, few studies have examined the effects of response cards on the vocabulary retention. Therefore, the current study focuses on the use of write-on response cards applied to vocabulary acquisition task.

**Vocabulary**

Before applying response cards to vocabulary acquisition, it is important to operationalize what types of skills constitutes “vocabulary” and its acquisition. This section provides an overview of studies of vocabulary development and vocabulary instruction.

*Development of Vocabulary and Relation to Other Skills*

Formation of an individual’s vocabulary knowledge starts from a very early stage of one’s life through exposure to the spoken language of adults such as primary caregivers (Hart & Risley, 1995) or peers (Wise, Sevcik, Morris, Lovett, & Wolf, 2007), or through exposure to stories (Silverman, 2007). In their report of a longitudinal study, Hart and Risley (1995) described how children’s frequency and amount of exposure to different forms of words (i.e., lexicon) impacted the rate of vocabulary acquisition and predicted later performance in reading comprehension or other test performance at third grade. In addition, they described how the gap between children with low vocabulary knowledge and high vocabulary knowledge continues to exist and widen as children from both groups proceed through the school system. Typically, children learn their words through verbal interactions at home and come to school with thousands of words.
However, children from most supportive environments with ample language experiences learn far more words than children who are from environments with much less verbal communication (cf. Ebbers & Denton, 2008).

To illustrate the gap between students with and without vocabulary knowledge, other researchers have documented the negative impact of insufficient vocabulary knowledge on comprehension or other areas of learning. For example, Aarnoutse, Van Leeuwe, Voeten, and Oud (2001) conducted a longitudinal study examining how decoding efficiency, reading comprehension, vocabulary, and spelling develop over the 6 years of elementary education. They also examined the differences between poor, average, and good performers with regard to the development of decoding efficiency, reading comprehension, vocabulary, and spelling of three cohorts of students drawn from 39 elementary schools in the Netherlands. The results suggested that the difference among low-, medium-, and high-performance groups continued to exist although students in all groups benefited from instruction in all skill areas. Based on the results of these studies, importance of early exposure to language is one of the most important predictors of later reading and language achievement.

Cunningham and Stanovich (1997) also demonstrated how reading or other abilities demonstrated early can predict later academic performance. In their study, 56 first grade students (32 boys and 24 girls) were assessed for cognitive and reading ability. The Peabody Picture Vocabulary Test and the Raven’s Colored Progressive Matrices were administered as their cognitive measures, and the Metropolitan Achievement Test, MacGinitie Reading Tests, and Wide Range Achievement Test were administered as reading ability measures. Ten years later, 27 11th grade students (15 boys and 12 girls)
took the Nelson-Denny Reading Test, or the Peabody Picture Vocabulary Test. Further, Print Exposure measures were administered as reading comprehension, written and picture vocabulary, and print exposure measures. The results indicated that both cognitive and reading measures in first grade strongly predicted the 11th-grade outcomes. As these studies demonstrate, vocabulary development at early stage serves as a critical factor that predicts later academic achievement.

**Vocabulary Instruction**

For those children who may have limited early exposure to language, some form of intervention is needed to address the lack of vocabulary knowledge, especially when research has repeatedly suggested the importance of vocabulary knowledge as the basis for comprehension (e.g., Boulware-Gooden, Carreker, Thornhill, & Joshi, 2007; Lubliner & Smetana, 2005) and as ultimately leading to learning in many other academic areas such as language arts, science, and social studies. Naturally, acquisition and retention of vocabulary words are critical prerequisite skills to influence the success of subsequent learning (Coyne et al., 2007; Jitendra, Edwards, Sacks, & Jacobson, 2004). Many studies have accumulated evidence that suggest effectiveness of vocabulary instruction in improving students’ vocabulary knowledge and collateral improvement in other areas such as comprehension or word use.

According to these studies, vocabulary instruction for young students tends to employ more naturalistic activities, rather than more systematic and contrived academic activities (Coyne, Simmons, Kame’enui, & Stoolmiller, 2004). To illustrate the effectiveness of a naturalistic approach to vocabulary instruction, Braham and Lynch-Brown (2002) demonstrated that teachers’ read-aloud styles for storybooks
differentially affected students’ vocabulary and comprehension performances at the first and third grade levels. Teachers used just-reading (asking students to listen carefully to the story without asking questions or making comments), performance reading (asking questions about a purpose statement and targeted specific words and concepts for the total of 5 minutes), and the interaction reading (interspersing story-related questions and engaging in a discussion for 15 minutes). Following the read-aloud, students took tests comprised of 40 multiple-choice vocabulary and 17 multiple-choice comprehension questions. The results showed superiority of both performance reading and interactional reading formats over the just-reading format. Based on this study, we may be able to utilize story read-alouds in a systematic manner to promote development of vocabulary.

Other researchers have also used storybook reading time as an intervention for promoting vocabulary development. For example, Silverman (2007) used storybook reading and varied the intensity of additional explanation for three different instructional formats on kindergarten students’ performances vocabulary knowledge measures. One of the conditions used contextualized instruction, which involved the teacher leading a discussion about new words that appeared in the storybook and encouraging children to connect the words with their background knowledge and experiences. In another condition, analytical instruction was added to contextualized instruction. In analytical instruction, students were encouraged to analyze words by thinking about application of the words in various contexts. In the third condition, anchored instruction was added to contextualized instruction and analytical instruction. Anchored instruction involved sound and letter instruction. The results showed that allowing children to engage in analyses of word sound and meanings was more effective than just relating the words to
their background knowledge. Coyne, McCoach, and Kapp (2007) also used a storybook reading activity in a small-group context and compared extended vocabulary instruction with incidental exposure and embedded instruction on kindergarten students’ expressive and receptive definitions of words and performance on Peabody Picture Vocabulary Test-III. Extended instruction is characterized by explicit teaching of definitions and pronunciation of target words in varied contexts and formats. Embedded instruction also provides the definition of target words within the context of the story but provides fewer encounters with target words. The results indicated that the extended instruction, which allowed for more encounters with the target words, resulted in greater word learning and also maintained much of the understanding of word meaning for six to eight weeks.

As students become older, however, there is a marked increase in the number of words they encounter every year. Anderson, Wilson, and Fielding (1988) documented that average students in grades 3 through 12 learn approximately 3,000 words a year if they read between a half-million and a million words of running text. Due to this sheer number of words, older students, in contrast to young students, may need to focus more on academic content and experience large amounts of printed materials representing a variety of genres (Reis, Jacobs, McCoach, & Coyne, 2008) through a variety of instructional formats (McDaniel & Pressley, 1989). Available instructional strategies include (a) directly teaching vocabulary words (e.g., McKeown, Beck, Omanson, & Pople, 1985), (b) mnemonic keyword strategies (e.g., Levin, McCormick, Miller, Berry, & Pressley, 1982), (c) peer tutoring (Malone & McLaughlin, 1997), or (d) computer-assisted instruction (Boling, Martin, & Martin, 2002).
For example, McKeown and colleagues (1985) compared the differential effects of three types of instruction (rich, rich/extended, and traditional) under two varying levels of frequency with which fourth grade students encountered the target words during an instructional session (high = 12 encounters per word, low = 4 encounters per word) on students’ performances on definitional knowledge, fluency of access to word meanings, context interpretation, and story comprehension. The results indicated that both rich and extended instructional methods were equally effective in improving all measures except for story comprehension, which showed higher performance with the rich/extended instruction that included multi-component definitional and contextual vocabulary instruction in and outside of the classroom. Also, high exposure to target words produced better results for all measures. In other cases, mnemonic strategy instruction was used to teach definitions of words to fourth grade students (Levin et al., 1982) and eighth grade students (Levin, Levin, Glasman, & Nordwall, 1992). In the Levin et al. (1982) study, fourth grade students were asked to verbally provide definitions of words instructed using mnemonics (keywords) or their own study methods. The results suggested the superiority of the mnemonic strategy approach when the measures involved giving the definitions of the target words. Levin and colleagues (1992) examined the comparative effects of mnemonic strategy instruction and the free study strategy of eighth grade students on their performances of definition recall and sentence comprehension measures. The results indicated that mnemonic keyword strategy produced superior performances on both measures.
Malone and McLaughlin (1997) examined the effects of reciprocal peer-tutoring in conjunction with a group contingency (RPT+GC) on acquisition of definitions of vocabulary to seventh and eighth grade typically developing students. In an ABAB and BABA time-series design, the students’ performances on the weekly definition tests of target vocabulary in the RPT+GC condition was compared to those of the traditional instructional design. The results indicated that students performed better on the weekly definition test in the RPT+GC condition.

Despite much research on vocabulary instruction, there are several applied questions that remain unanswered. These points concern the procedural issues associated with the studies. First, the optimal instructional environment or procedures have not been determined. Vocabulary instruction sometimes takes place in the context of story reading for elementary students (e.g., Brett, Rothlein, & Hurley, 1996) or for kindergarteners (e.g., Coyne et al., 2007). In other cases, vocabulary was taught in collaboration with speech/language therapy and classroom instruction (e.g., Parsons, Law, & Gascoigne, 2005) or in elementary-level classrooms (e.g., Apthorp, 2005; Nash & Snowling, 2006), or in clinic setting (e.g., Fossett & Mirenda, 2006). Second, measures used to assess students’ “vocabulary knowledge” are also inconsistent in terms of form and the levels of processing required with the task. Those measures included combinations of simple identification tasks such as (a) multiple choice tests (e.g., Braham & Lynch-Brown, 2002; Brett et al., 1996), (b) sight word reading, (c) identification of examples, (d) tasks that require “higher-order thinking” such as oral recall of vocabulary definition (e.g., McDaniel & Pressley), (e) production of examples of vocabulary words (Fossett & Mirenda, 2006), or (f) standardized assessment batteries (e.g., APTHORP, 2006). As these
studies show, choosing measures that truly illustrate learners’ vocabulary knowledge is a complex and difficult task (Parsons et al., 2005); it requires a series of determinations at multiple levels, such as the mode of target response, dimension of language, and complexity of thinking. In addition to the aforementioned two points, the number of studies that employed single-subject designs to teach vocabulary to students with disabilities is limited. The next section will provide some insights about the reason for the perceived lack of consensus as to how vocabulary should be taught and measured (NRP, 2000).

*Dimensions of Vocabulary Knowledge*

One reason for the lack of consensus regarding how vocabulary knowledge is measured is its complexity. Vocabulary knowledge is considered to consist of different levels of knowledge. Language has traditionally been described in terms of five dimensions concerning form (phonology, morphology, and syntax), content (semantics), and use (pragmatics) (Heward, 2006). Phonology refers to the linguistic rules governing a language’s sound system. Morphology refers to the study of the basic units of meaning and how these units are combined into words. Syntax concerns the rules that govern the meaningful arrangement of words into sentences. Semantics is the meaning of words and relationships between words. Pragmatics refers to a set of rules regarding how language is used in communication (Heward, 2006; Nation, Snowling, & Clarke, 2007). Vocabulary knowledge involves how the sounds are connected to form words (phonology), how parts of words are connected (morphology), meaning of words (semantics), and how words are used (pragmatics). In order to truly illustrate the complexity of vocabulary knowledge, syntax, semantics and pragmatics should be given
special consideration. A word usually has multiple meanings (semantics), and our interpretation depends to a great deal on the context in which the word is used (pragmatics). Also, depending on the meaning and how it is captured in the context, the same word can be treated as a noun or a verb, requiring a different sequence of words within a sentence (syntax). For example, the word “rail” has multiple meanings as a noun, and also serves as a verb with a different meaning. Especially when the word is used as a noun, a reader has to rely on the context to interpret its true meaning. In addition, vocabulary knowledge sometimes involves spelling, or relationship between sounds (phonemes) and their written forms (graphemes).

Some researchers have focused on these dimensions of vocabulary and used one of them to teach other dimensions of vocabulary. Baumann, Edwards, Boland, Olejnik, and Kame’enui (2003) compared morphemic and contextual analysis instruction (MC) with textbook vocabulary instruction (TV) to teach social studies vocabulary to 8 fifth grade students. Students took textbook vocabulary tests, immediate vocabulary-in-context tests, comprehension tests, chapter tests, and delayed vocabulary-in-context tests. The results of this study indicated that TV students learned more textbook words, MC students inferred more meanings of novel affixed words, MC students were more successful at inferring the meanings of morphologically and contextually decipherable words on a delayed test but not on an immediate test, and there was no difference between the groups on a comprehension measure or a social studies learning measure. This study indicated that teaching one of the dimensions of vocabulary as a strategy can enhance generalized performance on untaught items.
Due to the complex nature of vocabulary knowledge, measurement of vocabulary knowledge has to assess these different dimensions. To establish a new word in the lexicon, a student needs to analyze the available phonological, semantic, and syntactic information of the word and map one piece of information onto another (Parsons et al., 2005). Vocabulary knowledge can be conceptualized using two different frameworks that overlap with each other. The first concept concerns the mode of response, whether the vocabulary is orally uttered or written. When a reader encounters an unknown word in a text but can decode it phonetically, and if the decoded word is in the reader’s oral vocabulary, he or she can understand the meaning. Similarly, a reader who can use the contextual clue (print) to figure out the meaning of the unknown word has print vocabulary knowledge (NRP, 2000). Second, vocabulary knowledge can be conceptualized as to whether assigned tasks are receptive or expressive (Nash & Snowling, 2006). Examples of receptive tasks involve requiring children to make some physical responses (e.g., “Show me X” or circling a correct answer from possible options). Expressive tasks are exemplified by asking students to verbally tell association of an item (e.g., “Tell me as many animals as you can.”) or definition of words, or to write definitions of words or use the words in novel sentences. This second consideration pertains to the first consideration, mode of response, because expressive or receptive tasks can take both written and spoken forms. In their vocabulary instruction for third and fifth grade students, Nelson and Stage (2007) had the students to engage in a series of receptive and expressive tasks during “contextually-based multiple meaning vocabulary instruction.” For example, students were asked to write new sentences using the target word or its related word, or put the target word in a word map and match with appropriate
definition. Armbruster, Lehr, and Osborn (2003) suggested that students’ understanding of the meanings for a word is enhanced when they are asked to generate contextually-correct written narratives using its meanings.

Summarizing the implications from the available research on vocabulary instruction, we can draw some insights regarding the components of effective vocabulary instruction. Vocabulary instruction should involve the following: (a) a definition and explanation of word meanings (Arthorp, 2006; Nash & Snowling, 2006), (b) examples of words used in context, (c) extended discussion of the use of the target words and related words (Nelson & Stage, 2007), and (d) both receptive and expressive activities (Wise et al., 2007). However, complexity of vocabulary coupled with the large amount of words makes execution of all these factors difficult. Therefore, an important consideration is how we can achieve effectiveness and efficiency simultaneously (Jenkins & Dixon, 1983).

**Curriculum-based Measurement (CBM)**

Lack of consistency in measuring vocabulary performance presents a serious concern for the generality of available studies. More specifically, without consistent means of reporting student performance trends across studies, we cannot make meaningful comparisons. Research in vocabulary instruction awaits standard and consistent measurement that can be used across studies and to assist in identifying deficiencies in specific areas of students’ academic functioning (Davis, Caros, Grossen, & Carnine, 2002). Standardized curriculum-based measurement (CBM) was developed by the initiative of special education in response to limitations inherent in commercially available assessment tools or teacher-developed informal measurement tools (Deno,
CBM is characterized by measures that are sensitive to the little changes in students’ performances, repeated measurement, relative ease of implementation, and availability of standardized data (Hosp & Hosp, 2003). Appropriate use of CBM can enhance teachers’ abilities to monitor progress of student performance and to adjust their instruction accordingly (Capizzi & Fuchs, 2005; Madelaine & Wheldall, 2005). Teachers who use CBM can also enhance the effectiveness of communication among other professionals because of the standardized protocol for assessment and interpretation (Deno, 1985). CBMs are used in the areas of reading, spelling, math (Hosp & Hosp, 2003), and in writing (Espin, Weissenburger, & Benson, 2004).

These characteristics of CBMs can address the difficulty associated with identifying valid and efficient measures for vocabulary knowledge, especially for assessment of students’ expressive vocabulary in the written form: spelling and writing. Part of the vocabulary knowledge is assessed by asking students to spell the target words or synonyms of these target words. For this type of assessment, the spelling CBM provides a stringent measure of students’ ability. Although the formal spelling CBM is conducted by having the students spell words from dictation for 2 minutes and by counting the number of correct letter sequence (CLS) (Hosp & Hosp, 2003; Shinn & Shinn, 2002), the current study did not involve 2-minute dictation but used the CLS as a more stringent measure of students’ spelling performances within an experimenter-developed assessment. A correct letter sequence (CLS) is analyzed for pairs of letters within a word; if a sequence of two letters is correct, it is indicated using a caret (^). In addition, the spaces prior to the first letter and after the last letter are also counted as a CLS if these letters were written correctly. If only one of the letters in the two letter
sequence is correct, no credit is given. Similarly if the first or last letter is incorrect, no credit is given for that sequence. In these cases, a minus (-) is placed to indicated an incorrect letter sequence (Hosp & Hosp, 2003). Examples of CLS or CLS involving an incorrect letter sequence are presented below, using the word “book.”

\[^b^o^o^k^\] (total of 5 CLS; correct spelling)

\[^b^o^-k^\] (total of 3 CLS; incorrect spelling)

\[-d^-o^-c^-] (total of 1 CLS; incorrect spelling)

\[-d^-u^-c^-] (total of 0 CLS; incorrect spelling)

As the examples above illustrate, CLS is more sensitive than determining correct/incorrect spelling by examining the word as a whole. Although examples 2 through 4 are incorrect, a student who writes “bok” can be considered as having a better phonemic awareness or a greater understanding of alphabetic principle than another student who writes “book” as “dooc” or “duc.” With associated ease of implementation and standardized nature, employing CBM as a part of vocabulary instruction benefits both students and teachers. Teachers can easily implement the assessment and obtain valid information to adjust their instruction; students also benefit from the efficient instruction provided by the teachers.

In this study, however, I adapted the CLS so that I could utilize its sensitivity to capture small changes in the participants’ performances for the purpose of analysis. With this adaptation, two characteristics of CBM were omitted in this study. First, CLS was not used to monitor progress and inform instruction. Second, although CBM is usually implemented within a standard unit of time (e.g., 17 words dictated/spelled in 2 minutes), I did not include this component during the vocabulary probe.
Massed Versus Distributed Practice

Along with the consideration of how we should measure student performance on vocabulary tasks, a consensus should also be built regarding how to design instruction, especially when vocabulary overlaps with different skill areas and encompasses a wide range of skills. Accumulated empirical evidence indicates that acquisition of a new skill and its retention are enhanced when students receive ample practice opportunities and frequent feedback (Cull, 2000). Crawford and Baine (1992) suggest that rate of acquisition and length of maintenance are positively affected by repeated practice and present three variations with which teachers can design repeated practice opportunities for students: distributed practice, spaced practice, and massed practice. Distributed practice refers to repetition of a task $T$ with one or more other tasks in between each trial of $T$; with spaced practice, each trial of task $T$ is interspersed with rest periods; massed practice refers to presentation of task $T$ without intervening alternative task trials or rest periods between successive trials (Crawford & Baine, 1992).

Practice Formats

There are many studies that examined the comparative effects of these different practice formats. Cepeda, Pashler, Vul, and Wixted (2006) reviewed 317 experiments located in 184 articles. Effects of spacing (consecutive massed presentations vs. spaced learning episodes) and lag (less spaced vs. more spaced learning episodes) were examined, as well as expanding interstudy interval (ISI). The analysis suggested that, although inconsistent across studies, the longer the interval between study episodes, generally the longer the retention. One factor that accounts for the inconsistency in the available results resides with inconsistency regarding how these practice formats are
operationalized and used, especially with applied studies (Crawford & Baine, 1992).

Originally, effects of distributed practice and massed practices were studied in the basic animal study context. Fanselow and Tighe (1988), for example, used rats’ postshock freezing response as a measure and observed for the rate in which 36 rats exhibited postshock freezing in the same chamber where they have received three electric shocks in three schedules that differed in the inter-shock intervals. In this study, the postshock freezing was “learned” and was maintained better for the rats under a more distributed schedule; that is, those rats that received shocks with longer intervals tended to exhibit more responses 24 hours later.

When the scheduling effect was applied to human behavior in a simple paired-associate task, the same time-based paradigm was used with lists of words (e.g., Forrester, 1969). The results were not always consistent with those of the basic animal study. When examining both cases, the difference between the animal study and the human study is what is interspersed between each leaning trial. With animal studies, only one stimulus was taught to the subjects. With human studies, in addition to varying the length of intervals, different words were used. Therefore, distributed and massed practices with human studies differed in two aspects: length of inter-trial intervals and the stimuli (i.e., things children learned). For humans who need to learn complex behaviors, practice schedules should be conceptualized and examined with three different formats: massed, distributed, and spaced practices. Also, tasks such as paired-associate are arbitrary and nonfunctional; therefore, the paradigm should be examined with more functional tasks.
Confusion of the terminology and operationalization of the practice schedules continue to exist when applied to functional skills of humans (e.g., Rider & Abdulahad, 1991), but some attempts have been made to address this concern. Using an alternating treatment design, Mulligan, Lacy, and Guess (1982) examined the efficacy of massed, distributed, and spaced practices on the accuracy of four different motor and/or cognitive skills (e.g., setting table, picture/object matching, folding, and manding) of 11 students with severe disabilities. In the massed practice condition, all the training trials of one program were completed before the trials for the next program began. Sessions in the spaced condition followed the same type of sequence, but there was a 15 to 20 seconds between any two trials. Sessions in the distributed practice condition ordered the trials so that a cluster of four trials, one from each training program, was repeated 10 times. The results indicated that distributed practice created a significant difference in the accuracy of responding and rate of refusals to participate in the study. This study was noteworthy in the way it operationalized how massed, distributed, and spaced practices were applied to functional responses of individuals with severe disabilities.

Along the same line, Dunlap (1984) compared distributed and massed practices to teach several tasks involving recognition, discrimination, and sequencing to 5 children with autism ages 4 to 10 years old. In the massed practice condition, only one acquisition task was presented. There were two variations for the distributed practice condition: varied-acquisition-task, in which 10 acquisition tasks were randomly interspersed throughout each session; and varied-with-maintenance-task, which involved random interspersal of 5 acquisition and 5 previously-mastered tasks. Results indicated that not
only was the distributed practice more effective in acquisition of skills, but also that the varied-with-maintenance-task practice resulted in greater retention than the varied-acquisition-task.

In addition to the motor or recognition skills, the massed and distributed practice paradigm was applied to academic behaviors of students with different characteristics. Bloom and Shuell (1981) compared the effects of massed and “distributed” practice of French vocabulary words of 56 students enrolled in a secondary-level French course in a high school. However, the so-called “distributed practice” in this study was a combination of spaced practice and distributed practice of the same 20 words. Students in the distributed practice condition had 10 minutes to practice their words for three days; each day, a different activity was administered (i.e., multiple-choice, written fill-in-the-blank, and written spelling tests). Students in the massed practice condition studied all 20 words using all three activities in a day for total of 30 minutes. Immediately following the last study period, students took a vocabulary test which consisted of the multiple-choice, fill-in-the-blank, and spelling prompts. Also, students took the same test seven days later without additional warning or studying. The results showed that performances on the immediate test did not differ across groups, but distributed practice produced significantly superior performance. However, since the distributed condition in this study combined both skill (activity) distribution and time spacing, perceived superiority of this condition can only be interpreted as the function of repeated exposure to the target words. To address this concern, the massed practice condition should have been arranged so that the spacing effect could be eliminated.
Gettinger, Bryant, and Fayne (1982) used distributed skill practice to teach spelling to 39 elementary students with learning disabilities. Students in the experimental group received instruction repeatedly, and the control group received the classroom teacher’s regularly implemented instruction. The results suggested that the distributed practice of spelling was more effective in improving students’ accuracy of spelling and the accuracy of transfer word spelling (generalization). However, both the distributed practice group and the teacher instruction group received a series of instructions for the three-week-period, involving both spaced and possible distributed skill practices. Coupled with the fact that the targeted skill was limited (i.e., spelling), further research is needed to make the conditions more comparable. Also, additional research should look at how massed and/or distributed practices affect the dimensions of vocabulary knowledge that have not been studied yet.

Aside from this line of study, there is an extensive amount of evidence accumulated for the efficiency and effectiveness of distributed and massed practices in the educational context. One application is found within the Direct Instruction (DI) program (Carnine, Silbert, & Kame'enui, 1990). In this program, content is first taught using intense, massed practice to initially establish a learned response. After a sufficient level of proficiency has been attained, instruction shifts into a “firming cycle,” where additional practice increases response fluency. Review sessions during this cycle are brief, and interspersed into instruction of new materials. DI, however, is more than just a package of reading, vocabulary, or math instruction; it encompasses effective teaching practices, curriculum design, progress monitoring, and classroom management (Stein, Carnine, & Dixon, 1998). The vocabulary instruction (i.e., the independent variables)
used in this study reflect all these factors. It includes a lot of ASR opportunities, systematic presentation of instructional content in massed and distributed formats, and effective progress monitoring using curriculum-based measurements (CBM). Since ASR strategies can diminish problem behaviors simultaneously with enhancing the academic performance, the response card strategy meets both effective teaching practice and classroom management requirements.

The available massed versus distributed practice studies show, massed and distributed practice studies have generally supported the relative effectiveness of distributed practice to foster maintenance of acquired skills. However, these practice formats have been examined with very specific skills such as retention of spelling. Effects of these practice formats on “vocabulary” as a whole have not been directly studied, except for in the context of DI curriculum. However, vocabulary in DI curriculum is embedded as part of the complex construct of “reading”; therefore, it is difficult to separate out the effects of massed and distributed practices on development of vocabulary knowledge out of other reading skills. Additional studies are needed to examine the specific effects of massed or distributed practices on different dimensions of vocabulary acquisition, maintenance, and generalization.

**Summary**

The existing literature on vocabulary instruction has revealed some features of effective vocabulary instruction. First, there should be guided practice with timely feedback, implemented in conjunction with initial massed practice and gradual shift to distributed practice. Second, there should be active student engagement embedded throughout the instruction. The response card strategy used in this study can achieve both
of these features by providing ample response opportunities for students and allowing teachers to monitor students’ responses in an on-going manner. Also, massed and distributed practices will be used in the context of content practice, instead of the spaced practice.
CHAPTER 3

METHOD

In the following sections, I will describe the method used in this study. Specifically, I will provide detailed descriptions of the setting, the participants, materials, independent variable, and dependent variables, as well as the instructional and data-collection procedures.

Setting

This study took place in a seventh grade language arts classroom in a large suburban middle school that served students from grades 6 through 8. In the language arts classroom, there were a total of 21 students, a language arts teacher, and two assistants for students with disabilities. These assistants were generally seated next to the two students who received special education services; however, during the instruction for this study, they did not provide any assistance to the target students and stayed away from the students with special needs. Until the end of week 4, there was also a student teacher who taught lessons with the language arts teacher. The physical layout of the classroom included desks and chairs assembled in rows facing the whiteboard at the front of the room. In addition, the classroom was equipped with an LCD projector and three computers on the side of the classroom.
Participants

Students

Five students in the seventh grade language arts class were selected as target students in this study. These students were selected from the total of 21 students recruited to participate (Appendix A). Three of the students were selected because they receive special education services under the categories of mild mental retardation, speech and language impairments, or other health impairment. In addition, two students without disabilities were identified as comparison students. These comparison students were selected because (a) their parents had provided consent for their participation, (b) they had regular attendance in the language arts class, and (c) they were identified by the language arts teacher as performing well on a regular basis.

Student 1 (Jessie). Jessie received special education services under the category of mild mental retardation. Her chronological age was 12-9 at the beginning of the study. Her IQ score obtained in February 2007 using The Wechsler Intelligence Scale for Children Fourth edition (WISC-IV) indicated an overall IQ of low average to average. Her basic reading skills appeared to be on grade level as of March 2007. More specifically, she read 119 words per minute from a seventh grade text. On the reading section of the Ohio Achievement Test, she scored at the proficient level (score of 410 with 400 considered proficient). Her vocabulary scores were not available, but she did have a vocabulary goal on her IEP. As for her writing, the language arts teacher indicated in her IEP that “current weaknesses (of writing included) editing and putting ideas into paragraphs.” Based on these needs, she received special education services for written expression for one period per day with the intervention specialist coming into the general
education classroom and reading comprehension and vocabulary instruction five times a week in the resource room. She usually received a B+ in the vocabulary and writing activities assigned by the language arts teacher.

**Student 2 (David).** David received special education services under the category of speech and language disabilities. His chronological age was 12-8 at the beginning of study. His intelligence score obtained in June 2007 using *The Wechsler Intelligence Scale for Children Fourth edition* (WISC-IV) indicated an overall IQ of borderline to low average and low average to average, respectively. His oral reading fluency rate was 75 words per minute, which placed him at the 25th percentile in sixth grade. His vocabulary scores were unavailable, but the language arts teacher indicated writing as one of his weaknesses and commented on “the lack of details, grammar, legibility of writing and editing.” In a writing curriculum-based measurement (CBM) probe administered in June 2007, he wrote 45 words in a three-minute free writing. Based on these needs, he received special education services in written expression for one period per day with the intervention specialist coming into the general education classroom. He usually received a C+ in the vocabulary and writing activities assigned by the language arts teacher.

**Student 3 (Rachel).** Rachel received special education services under the category of other health impairment (OHI). Her chronological age was 13-0 at the beginning of the study. She was a relatively new student, so her recent intelligence scores were not available. However, a *The Wechsler Intelligence Scale for Children Fourth edition* (WISC-IV) administered in March 2006 indicated an overall IQ of average to low-average. She scored at the proficient level on the sixth grade Ohio Achievement Test (OAT). She scored 4/4 on the fifth grade level on the spelling portion of the *Brigance*
Inventory. As for writing, the language arts teacher indicated that her writing was at the basic level. Based on these needs, she received special education services for written expression for one period a day with the intervention specialist coming into the general education classroom. She usually received an A- on the vocabulary and writing activities assigned by the language arts teacher.

**Student 4 (Kelly).** Kelly was a comparison student who was identified by the language arts teacher as having good vocabulary knowledge. Her chronological age was 12-6 at the beginning of the study and she usually received grades of A- from the language arts teacher.

**Student 5 (Sam).** Sam was also a comparison student who was identified by the language arts teacher as having good vocabulary knowledge. His chronological age was 13-3 at the beginning of the study and he usually received grades of A- from the language arts teacher.

**Language arts teacher.** The language arts teacher had 12 years of experience teaching language arts at the middle school level. She received a Bachelor’s degree in elementary education from Northern Illinois University and a Master’s degree in children’s literature from The Ohio State University. She selected curriculum activities, developed lessons, and delivered them during the regularly scheduled class time. During the intervention for this study, she approved the vocabulary words and example sentences selected by the first experimenter as appropriate for her curriculum and ability levels of her students.

**Experimenter**

The experimenter was a doctoral candidate in special education and applied
behavior analysis at The Ohio State University. She was responsible for selecting target vocabulary words of the week, developing instructional materials, implementing lessons, collecting performance data, and implementing questionnaires to obtain students’ and teacher’s opinions about the two formats of vocabulary instruction. She received her M.A. in special education and applied behavior analysis from The Ohio State University in 2004 and had experience working as a part-time teacher for five years. She also had a year of experience working as a behavioral therapist for young children with autism.

Observer

Another doctoral student in special education and applied behavior analysis served as an observer who was in charge of collecting the reliability and procedural integrity data. At the beginning of the study, he was a second year doctoral student. He had 6 years of experience working with children with developmental disabilities in both home and school settings delivering behavioral interventions. He received his Bachelor’s degree in psychology and a Master’s degree in psychology with an emphasis in applied behavior analysis from California State University, Los Angeles.

Independent Variables

The independent variable used in this study was two practice formats embedded in a response card activity: massed and distributed practices of four aspects of vocabulary knowledge (i.e., spelling, definition, parts of speech, and synonym). In the massed practice condition, students practiced one aspect of vocabulary in a chunk for five different words of the day. More specifically, massed practice of the four aspects of vocabulary knowledge consisted of spelling practice of five words, followed by definition, parts of speech, and synonym practices of the same five target words using a write-on
The distributed practice (DP) condition was designed so that students practiced spelling, definitions, parts of speech, and synonyms of one of the words and then move to the four aspects of vocabulary for the second, third, forth, and the fifth word. Both distributed and massed vocabulary practices were embedded in a write-on response card activity in order to ensure active student response during the instruction. Upon prompts from the experimenter, students wrote on their write-on response cards the spelling, part of the definition, and a letter from a multiple-choice prompt indicating the appropriate example or synonym and the part of speech.

Selection of Words and Materials

Selection of Vocabulary Words

The experimenter developed two lists of words that were used for instruction in either the distributed or massed practice format. The target words were taken from the Scholastic Scope Teacher’s Edition (2004; 2007) and SAT vocabulary lists from the Princeton Review (2007) and Kaplan (2005), excluding words that matched or shared the root with the 80 words that had previously been taught during the school year in the language arts class. For example, the word “consistently” was excluded from the list because the students had already learned the word “consistent.” Also, the first experimenter consulted the language arts teacher and excluded three words that were indicated as familiar to the students (e.g., “environment”). These processes generated a pool of 301 words that were hypothesized as unknown to the students. Then, the 301 untaught words were sequenced from least to most in terms of number of letters contained in each word. This measure was taken to ensure the number of letters taught in the two instructional formats (i.e., MP and DP formats) would be identical. Words that
<table>
<thead>
<tr>
<th>Massed practice (MP) Condition</th>
<th>Distributed practice (DP) Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>banal (1)</td>
<td>rail (2)</td>
</tr>
<tr>
<td>insolent (1)</td>
<td>erudite (2)</td>
</tr>
<tr>
<td>fledgling (1)</td>
<td>adversary (2)</td>
</tr>
<tr>
<td>deft (1)</td>
<td>lofty (2)</td>
</tr>
<tr>
<td>lavish (1)</td>
<td>allude (2)</td>
</tr>
<tr>
<td>tyro (3)</td>
<td>laud (4)</td>
</tr>
<tr>
<td>cryptic (3)</td>
<td>taciturn (4)</td>
</tr>
<tr>
<td>conundrum (3)</td>
<td>sagacious (4)</td>
</tr>
<tr>
<td>lithe (3)</td>
<td>cynic (4)</td>
</tr>
<tr>
<td>myriad (3)</td>
<td>cajole (4)</td>
</tr>
<tr>
<td>lurk (5)</td>
<td>glib (6)</td>
</tr>
<tr>
<td>rebuke (5)</td>
<td>decorum (6)</td>
</tr>
<tr>
<td>convivial (5)</td>
<td>brusque (6)</td>
</tr>
<tr>
<td>prosaic (5)</td>
<td>aesthetic (6)</td>
</tr>
<tr>
<td>encroach (5)</td>
<td>debunk (6)</td>
</tr>
<tr>
<td>fetid (7)</td>
<td>ruse (8)</td>
</tr>
<tr>
<td>shrewd (7)</td>
<td>defile (8)</td>
</tr>
<tr>
<td>polemic (7)</td>
<td>thwart (8)</td>
</tr>
<tr>
<td>hackneyed (7)</td>
<td>adulation (8)</td>
</tr>
<tr>
<td>windfall (7)</td>
<td>decipher (8)</td>
</tr>
</tbody>
</table>

Table 3.1: List of Vocabulary Words for Each Condition

Note: Number in the parentheses indicates the session in which the word was taught.
consisted of 4 to 9 letters were selected as target words because there were a sufficient number of 4- to 9-letter words to run eight instructional sessions. Then, the words that contained 4 to 9 letters were divided into two groups, each representing the two instructional formats. Words with similar meanings (e.g., “rail” and “reproach”) were put into different groups so that each list contained the same number of synonyms. This process generated two lists of 85 words. Those words are listed in Table 3-1 by the week in which they were instructed.

**Instructional Materials**

Each write-on response card was made of a piece of letter-sized paper laminated with a letter-sized heat seal pouch. During instruction, each student was given a response card and an Expo® dry erase marker. The experimenter selected bold colors such as black, blue, green, or purple so students’ writings on their response cards were easily visible. Also, a piece of paper towel was given to each student so it could be used as an eraser. For the assessment of four different aspects of vocabulary knowledge (i.e., spelling, definition, parts of speech, and synonym), four separate letter-sized sheets that contained the quiz prompts were used. These four sheets were in different colors each session. At the beginning of each session, all of these materials were put in a letter-sized manila envelope named for each student and distributed to them. After the instruction, students were instructed to put all the materials back in the original envelope.

The week’s five target words were projected on a large whiteboard in front of the room through a power-point presentation so all students could see them. For the presentation, the experimenter used a laptop computer, an LCD projector, and a wireless presentation device so she could operate the instructional slides away from the computer.
**thwart** (verb)
means to prevent from doing something.
Another word for thwart is **hinder**.

(a) Model slide

(b) Spelling prompt slide

Note: Underlined word in the spelling prompt slide is what students should write on the response card

*Figure 3.1. Instructional slides.*  
(Continued)
2-2. Write definition

**thwart**
means to prevent from doing something.

2-3. Choose another word and part of speech

**thwart**

- play
- hinder
- wait

This word is

1. Verb
2. Noun
3. Adj.
while monitoring the performance of the students. Instructional slides were prepared by
the instructor each week. Time needed for preparation ranged from 15 to 45 minutes and
decreased as the study proceeded.

A week’s PowerPoint presentation consisted of a total of 20 slides. Four slides,
each of which represented four vocabulary aspects, were devoted to each word of the five
target words. These slides were (a) a “model” slide, which contained the word, its
meaning, the example of a synonym (e.g., “thwart means to prevent from doing
something. Another word for thwart is hinder.”) and appropriate part of speech (e.g.,
adjective); (b) a spelling prompt slide, which contained the blank for the word and
(corresponding definition (e.g., “_______ means to prevent from doing something.”); (c)
a definition prompt slide, which contained the word with the definition left blank (e.g.,
“thwart means ______________.”); and (d) a synonym and parts-of-speech prompt slide,
which contained the word and three choices from which students were to write the letter
(corresponding to the correct answer (e.g., “thwart A. play, B. hinder, C. wait”, “This
word is 1. noun, 2. adjective, 3. verb) (Figure 3-1).

**Dependent Measures**

The dependent variables in this study were grouped into five major categories:
vocabulary acquisition measures, spelling acquisition measures, generalization measures,
maintenance measures, and social validity measures. Vocabulary acquisition was assessed
though (a) accuracy of definition writing, (b) accuracy of identification of parts of speech,
and (c) accuracy of recognition of synonyms illustrating the vocabulary words. Spelling
acquisition was assessed through (a) the number of correctly spelled words and (b) the
percentage of correct letter sequences (CLS). Generalization was assessed through the
four-point ratings of the appropriateness of the target word usage in the novel sentences. As the maintenance measure, the vocabulary, spelling, and generalization measures were repeated two or three days after the instruction. Lastly, social validity of distributed versus massed practices was assessed through (a) the comparison of above dependent measures between the three students with disabilities and the two students without disabilities, (b) the ratings of distributed or massed practice formats by seventh grade students with and without disabilities regarding their vocabulary performances, and (c) the ratings of distributed or massed practice formats by the classroom teacher of seventh grade students with and without disabilities regarding the usability of the intervention procedures. Each dependent variable will be discussed in detail in the following sections.

**Vocabulary Acquisition Measures**

*Accuracy of definition writing.* The vocabulary probe contained five prompts to write the definition of the week’s five target words (Appendix B). A student’s definition writing was considered correct if it matched the original definition sentence or synonyms presented in the instruction slide or if it conveyed the same meaning as the original sentence/synonym presented during instruction (e.g., using “go down” instead of “decline”), and if it was grammatically correct. It did not have to start with a capital letter in order to be considered correct. Likewise, it could be correct even if the sentence contained grammatical errors or spelling errors as long as the sentence conveyed the original semantic meaning or the spelling errors represented the phonetic characteristics of the original sentence or word (e.g., “smely” for “smelly”).
The accuracy of identification of parts of speech. The vocabulary probe contained five multiple-choice prompts to which students circled the correct part of speech for each of the five vocabulary words (Appendix C). A response was considered correct if it matched the item in the experimenter-prepared answer key in the appropriate space.

The accuracy of recognition of synonyms illustrating the vocabulary words. The vocabulary probe contained five multiple-choice prompts to which students circled the letter that corresponded to the correct synonym for each of the five vocabulary words (Appendix C). A multiple-choice option for a word involved a correct option and two distracters. A response was considered correct if it matched the item in the experimenter-prepared answer key in the appropriate space.

Spelling Acquisition Measures

Total number of correct responses on a spelling probe. The vocabulary probe contained five spelling prompts to which the students wrote the spelling of the words as they heard the experimenter dictate each word (Appendix D). Students’ spelling was scored as correct if it matched the topographical characteristics of the same item in the answer key developed by the first author. A spelling response that did not match the answer key was still counted as correct if it started with a capital letter, contained only capital letters, or contained correct typographical variations (e.g., “a” instead of “o,” or “g” instead of “g”). A spelling response was counted as incorrect if it contained omission, substitution, insertion, or reversal of any letter(s).

Percent of correct letter sequences (CLS). This measure was used as a more sensitive measure of the students’ spelling performances. For each target word, the
number of correct letter sequences was counted and added together. In order to obtain CLS, each pair of letters within a word was analyzed whether the sequence was correct (^) or incorrect (−). For example, for the word “cat,” student’s responses would be analyzed as follows:

Student A’s response for the word “c a t”    ^c^a^t^  4 CLS
Student B’s response for the word “c a t”    −k−a^t^  2 CLS

If a target word involved a repetition of the same letter (e.g., “l” in “allude”) but had one of them omitted in a student’s response, the response was graded as follows:

Student C’s response for the word “a l l u d e”     ^a^l^u^d^e^  5 CLS

In this case, this student’s response would appear as involving no errors; however, this problem was solved by using the percent CLS.

In order to obtain the percent CLS, the number of CLS was divided by the total possible number of correct letter sequences. In this case, the denominator was the total possible number of CLS when no errors were made and the numerator was the number of CLS each student received. For example, the percent of CLS for words “cat” and “allude” was calculated as follows:

Student A’s “cat”    ^c^a^t^  4 CLS/4 Possible CLS = 100%
Student B’s “cat”    −k−a^t^  2 CLS/4 possible CLS = 50%
Student C’s “allude”    ^a^l^u^d^e^  6 CLS/7 possible CLS = 85.7%

Percent of CLS was used instead of number of CLS because the total number of possible CLS differed each week by two or three letters.
Generalization Measures

Accuracy of vocabulary words usage in novel sentences. The vocabulary probe contained five prompts to write five novel sentences (Appendix E) each containing one of the five target vocabulary words of the week. Each sentence was rated using a 4-point rating scale with 0 being the lowest and 3 being the highest. A rating of 0 was scored if (a) the student’s sentence was unfinished and/or did not contain the target word, (b) there was no response, or (c) the sentence did not make sense grammatically. For example, for the word “banal,” a 0 would have been scored for responses such as “My mother is very (unfinished)” or “My room is my banal.” A score of 1 was given for sentences that were grammatically correct but too general (e.g., “My story can be banal sometimes”) or if the target word was used in a different context and grammatically incorrect (e.g., “This food banal”). In this example, “banal” was defined as “commonplace” during instruction but used as “not fresh food” in the student’s sentence and grammatically incorrect. A score of 2 was given for sentences that were grammatically incorrect but used the target word correctly (e.g., “My speech banal when I don’t prepare well.”), or for sentences that were grammatically correct but used the target word in a different context (e.g., “This food is banal”). A score of 3 was given for sentences that used the target word correctly and were grammatically correct (e.g., “My writing can be banal when I don’t prepare well.”).

Maintenance Measures

The three vocabulary acquisition measures, the two spelling acquisition measures, and the generalization measure were repeated using the same vocabulary probe two or three days after instruction. Students’ responses were scored and analyzed in the same manner as the post-instruction probes were scored and analyzed.
**Social Validity**

Students’ and the classroom teacher’s opinions about the two instructional formats were solicited through questionnaires. Students were asked about the effects of distributed practice and massed practice of vocabulary words embedded in a response card activity (Appendix F). The classroom teacher was asked about the ease of planning, implementation, and evaluation of the two vocabulary practice formats (Appendix G). Also, in order to assess the social validity of the effects, the scores of students with disabilities were compared to those of students without disabilities who were identified by the classroom teacher as performing well on a regular basis (Student 4: Kelly and Student 5: Sam).

**Experimental Design**

An alternating treatment design (Cooper et al., 2007) was used to compare the effects of distributed versus massed vocabulary content practices embedded in a response card activity on the vocabulary knowledge of seventh grade students with and without disabilities. The distributed practices (DP) and the massed practices (MP) were alternated on a weekly basis.

An alternating treatment design was employed for the following reasons. First, the primary purpose of this current study was to compare the effectiveness of two instructional procedures that were basically similar to each other except for the sequence in which the vocabulary aspects were presented; second, the duration of the study was expected to be short due to the school events toward the end of the year. For these reasons, this design was chosen to be the most appropriate.
Data Collection Procedures

Probe Procedure

Immediately after each instructional session and three days after each instructional session, a 25-item vocabulary probe was administered to assess students’ knowledge on the five vocabulary words targeted for the week. However, in weeks 2 and 6, the maintenance probe was administered two days after the instruction due to snow day cancellation and school events, respectively.

The probe consisted of 4 separate sheets. The first page was devoted to spelling of the five words (Appendix D); the second sheet required students to write the definition of the vocabulary words of the week (Appendix B); the third sheet contained a synonym and two distracters for each of the week’s vocabulary words, as well as “verb,” “noun,” and “adj.” in parentheses so that students could circle the appropriate parts of speech for each word (Appendix C); the last sheet, which was printed on the back of the third sheet, prompted students to write five novel sentences each using one of the five target words of the week (Appendix E). Those four sheets were printed in different colors and the colors differed each probe session. Also, the probe sheets were presented in the order of spelling – definition writing – synonym/parts of speech – sentence writing (generalization) so that the probe sheets would not prompt correct responding for later sheets.

At the beginning of the probe, the experimenter instructed the students to take out the first sheet of the probe from the envelope by referring to the appropriate color (e.g., “Take out the pink sheet from the envelope”). Then, she instructed students to write their names at the top of the sheet. When all students complied, she started dictating the five target words, one every ten seconds. After dictating the fifth word, the experimenter
waited for 10 seconds and instructed students to put the first sheet back and take out the second sheet (e.g., “Now, put the pink sheet back in the envelope and take out the yellow sheet.”). After visually scanning the classroom and confirming that all students had taken out the second sheet, the experimenter instructed the students to write the definition next to each target word and allowed 40 seconds per word to write the definition (“Please write the definition of each word. You can write either the definition or synonym you practiced. You have 40 seconds for each word. Pencils ready, start.”). However, students progressively required less time to write the definition, so starting in week 3, the experimenter started to give 30 seconds per word to write the definition. When the time allocated for definition writing had elapsed, she instructed students to put the second sheet back in the envelope and take out the third sheet (e.g., “Put the yellow sheet back in the envelope and take out the blue sheet”). When all students complied, the experimenter said “Now, please circle the appropriate part of speech and a synonym for each word. After you are finished with the third sheet, turn the page over and work on another sheet. Please write a new sentence using each word. Please do not write the definition. I look forward to unique, interesting sentences. You have one minute for each word. Of course, you can move on your own pace. Please start.” The experimenter allowed 15 seconds for each word to identify synonym and parts of speech, and 1 minute to work on the sentence writing. However, students progressively required less time to work on the sheet. In weeks 3 to 5, students required the total of 5 minutes and 15 seconds. In week 6, students required 5 minutes and 11 seconds. In weeks 7 and 8, students required an average of 4 minutes and 59 seconds. When students were working on the probe, the experimenter walked around the classroom and monitored the students to ensure that they continued
working on the probe. Also, when she saw a student who had finished the last sheet, she instructed him/her to put the sheet back in the original envelope and collected it. In all probe sessions, the experimenter used a stopwatch so she could accurately implement the probe procedure. Also, she followed the probe procedure steps (Appendix I) to ensure correct implementation. After experimenter collected all students’ envelopes that contained their completed probe sheets, the experimenter thanked the students for their participation and left the room.

After the probe, the experimenter photocopied all the participants’ probe sheets and handed the copies to the second experimenter for the purpose of collecting reliability data. Both the experimenter and the observer used a data sheet (Appendix J) to record the target students’ performances.

**Instructional Procedure**

In a typical week, intervention consisted of three parts: instruction plus a response card practice activity, post-instruction probe, and maintenance probe. All students in the language arts class participated in all three parts. The instruction plus response card practice was a 10- to 12-minute choral responding and response card activity that involved a power-point presentation of the five target words and a sequence of instructional prompts. During the response card practice, each student used a write-on response card to respond to the experimenter-posed questions. After going through all the slides, students took a 10- to 13-minute probe in which they identified parts of speech and synonym and wrote the spelling, the definition, and a new sentence for each of the five words taught. Three (or two) days after the instruction, students took the same probe to see how they maintained the aspects of vocabulary learned in the instructional session.
Instruction (Introducing the words). On the instruction day, which typically happened on Mondays, five new words were taught across four different aspects in either distributed or massed practice format.

When presenting the model slides for the five target words, the experimenter introduced the content of the slide in a systematic manner using a choral responding technique (Heward, 1994). Specifically, when the target word appeared on the screen, the experimenter read the word out loud (e.g., “This word is “thwart”) and said “What’s the word?” so the students would respond back the pronunciation of the word (e.g., “thwart”) in unison. Next, the experimenter had the definition of the word appear on the screen, read it out loud (e.g., “This word means to prevent from doing something”), and asked “What’s the definition?” so the students would chorally read the definition sentence. Similarly, the experimenter read the part of speech (e.g., “This is an adjective”) and the synonym (e.g., “Another word for thwart is to hinder.”) out loud and had the students verbally repeat what she said. The number of opportunities for choral responding remained the same across both conditions. Lastly, the experimenter read a sentence that used the target word (e.g., “My friend accidentally gave away the plan for a surprise birthday party. She thwarted our plan”) (Appendix H).

Response card practice. During the response card practice, students practiced the vocabulary aspects (i.e., spelling, definition, identification of correct synonym, and identification of correct part of speech) in either massed or distributed practice formats (see Figure 3-2 for the prompt/practice sequence for each condition). When one of the prompt slides appeared during the presentation, the experimenter posed a prompt (e.g., “Write the word that comes in the blank” or “Write the letter of the correct example”),
waited for students to finish writing their responses, and gave the cue to hold their responses (e.g., “Boards up.”). After scanning the students’ responses, the teacher gave the appropriate consequence by praising the correct response, re-stating the correct answer, or repeating the same prompt so students would have repeated practice of the same content. When the experimenter finished presenting all the slides assigned for the day, she instructed the students to put the marker and the response cards away into the provided envelope and to take out the first sheet of the vocabulary probe.

In the MP condition, students practiced each vocabulary aspect (i.e., spelling, definition, identification of correct synonym, and identification of correct part of speech) in a chunk for all five target words. More specifically, the practice sequence went as follows: spelling (words 1 through 5), definition (words 1 through 5), identification of synonym and parts of speech (words 1 through 5) (Figure 3.2: massed practice). When students finished practicing all vocabulary components, the post-instruction probe started as described in the probe procedure.

In the DP condition, four aspects of a vocabulary word – spelling, writing definition, identification of correct synonym, and identification of correct part of speech – were presented in this order for each word during the response cards practice activity. More specifically, the practice sequence went as follows: spelling (word 1), definition (word 1), synonym (word 1), part of speech (word 1), spelling (word 2), definition (word 2), synonym (word 2), part of speech (word 2), spelling (word 3), definition (word 3), synonym (word 3), part of speech (word 3), spelling (word 4), definition (word 4),
Figure 3.2. Sequence of instructional prompt during the response card practice.
When finished practicing all vocabulary words, the post-instruction probe started as described in the probe procedure.

Procedures for Obtaining Reliability of Data

The second observer independently scored 50% of the students’ vocabulary probes. To do so, the experimenter photocopied students’ probe sheets for either post-instruction or maintenance each week. These unmarked photocopied vocabulary probes were used to collect the reliability data. IOA for all the dependent variables was calculated on the item-by-item basis to provide a conservative index of interobserver agreement (Cooper et al., 2007). More specifically, the IOA was calculated by dividing the total number of items with agreement by the total number of items and multiplying by 100. Before having the observer record the interobserver agreement (IOA) for all dependent variables, a short training session was conducted upon termination of the study. The experimenter gave the observer photocopies of three weeks’ worth of probe sheets completed by three non-participating students, data sheets (Appendix J), and answer keys. After scoring one student’s probe together, she asked the observer to independently score the remaining two students’ probe sheets. After the observer completed scoring, the experimenter and the observer together went through all the scoring and calculated IOA. These steps were continued until the IOA exceeded 90% for two students’ probe sheets from two different weeks. Agreement reached the 90% criterion after completing probe sheets from two different weeks.
Procedure for Obtaining Procedural Integrity

The observer observed the steps of vocabulary instruction for 25% of the sessions (i.e., two of the eight instructional sessions) using a checklist of the instructional steps (Appendix H). He observed one MP instructional session and one DP instructional session. The second experimenter placed a plus (+) for occurrence of each instructional step and a minus (-) for non-occurrences of a step. Percent of instructional procedural integrity was calculated by dividing the number of instructional steps implemented correctly within an instruction session by the total number of instructional steps (i.e., steps that were implemented correctly plus implemented incorrectly) and multiplying by 100 (Sterling-Turner, Watson, & Moore, 2002).

The observer also observed the steps of vocabulary probes for 12.5% of the probes (2 of all 16 probes) using a checklist of the probe steps (Appendix I). The second observer placed a plus (+) for occurrence of the probe steps and a minus (-) for non-occurrences of the probe steps. Percent of probe procedural integrity was calculated by dividing the number of probe steps implemented correctly within a probe session by the total number of probe steps (i.e., steps that were implemented correctly plus implemented incorrectly) and multiplying by 100 (Sterling-Turner et al., 2002).

Procedures for Obtaining Teacher’s and Students’ Opinions

Upon completion of the study, the students and the classroom teacher were asked to complete a questionnaire about the procedure and the effects of the two formats of vocabulary instruction. Students were required to complete the questionnaire
anonymously in order to avoid the interference of the experimenter knowing individual responses (Appendix F). The teacher’s questionnaire involved open-ended questions so that free opinions could be solicited (Appendix G).

Due to a school event (i.e., statewide achievement testing) that took place when the study was complete, the experimenter was not allowed to visit the classroom. Therefore, the language arts teacher gave the questionnaire to the students and had them complete them during study time. When they had finished the questionnaires, students submitted them to the language arts teacher, who, upon receiving all students’ questionnaires, put them in a large envelope and left them at the school office for the experimenter to pick up.

All but two students in the language arts class completed the questionnaire, including the five students who were identified as the participants to this study. Therefore, a total of 19 students completed the questionnaire. Given that these students completed the questionnaire anonymously and the experimenter was not present when they were completing the questionnaire, she could not separate the responses of the five target students from the other students’ completed questionnaires. For this reason, all available responses from 19 students were used as social validity data. All 19 students’ parents had signed the permission forms.

Students’ questionnaires contained eight questions regarding the purpose, method, and effects of the MP or the DP vocabulary intervention using response cards. In order to obtain their general opinions regarding the importance of learning new vocabulary (question 1), whether they liked using response cards (question 2), and the extent of helpfulness of response cards in learning vocabulary words were solicited using
a three-point Likert scale with 1 indicating “not at all” and 3 indicating “very much.” In order to solicit students’ opinions regarding the helpfulness of MP and DP, a four-point Likert scale was used with 1 indicating “not at all,” 2 indicating “a little,” 3 indicating “very much,” and 4 indicating “no difference.” To solicit students’ opinions regarding how much they liked either MP or the DP, a three-item multiple-choice question was used with “a” indicating the DP, “b” indicating the MP, and “c” indicating “no difference.” In order to solicit students’ opinions about the effects of the vocabulary intervention, the experimenter used four multiple choice items and asked the students to identify which of the vocabulary aspects that they thought had improved the most through the intervention. The last item asked the students to identify a practice condition they would like to use to learn vocabulary in the future using a three-point multiple-choice question, with “a” indicating the DP, “b” indicating the MP, and “c” indicating that they “don’t know.”

The teacher’s questionnaire contained questions about the procedures, results, and appropriateness of the intervention in the form of vocabulary instruction in the future. Also, her free opinions about the results were solicited.
CHAPTER 4

RESULTS

In this section, procedural integrity and the inter-observer agreement (IOA) data will be presented followed by each student’s performance on the dependent variables along with the students’ and teacher’s responses to items included in the social validity questionnaires. The five dependent variables will be discussed in separate sections. Accuracy of definition writing was measured by having the students write either the definition or the synonyms presented during the instruction. Semantically and phonetically identifiable responses were counted as correct even if students spelled the word incorrectly (e.g., “using “go down” instead of “decline” or “dicline”). Accuracy of identification of parts of speech and synonyms for five target words comprised one of the four sheets of the post-instruction and maintenance quizzes making up to the total of possible 10 points. However, these two measures will be analyzed separately. Accuracy of spelling was measured by having the students spell the dictated five target words. As a more sensitive measure of students’ spelling performances, percent of correct letter sequences was counted for the target words. Lastly, as a generalization measure, a four-point scale was used for the students’ novel sentences using each of the target words. In the following sections, the results are presented based on the research questions.
Believability of Data

Procedural Integrity

Instructional sessions. The observer came to two (25%) of the instructional sessions to see how well the experimenter followed the instructional sequences and the procedures using a checklist (Appendix H). The procedural integrity was 100% for both sessions.

Probe sessions. The observer came to two (12.5%) of the probe sessions to see how well the experimenter followed the probe administration sequences (Appendix I). The probe procedural integrity was 100% for both sessions.

Interobserver Agreement (IOA)

IOA was calculated separately for each dependent variable. For accuracy of spelling, IOA calculated by item-by-item agreement ranged from 80% to 100% with a mean of 96.4%. The disagreements were found on performances of students 1 (Jessie), 2 (David), and 3 (Rachel), whose writing tended to be illegible due to poor handwriting and/or repeated erasing. For percent correct letter sequences (percent CLS), IOA ranged from 88.2% to 100% with a mean of 97.7%. For accuracy of definition writing, IOA ranged from 60% to 100% with a range of 95.2%. The lower agreement for this dependent variable occurred with the two comparison students (i.e., Kelly and Sam) who tended to elaborate on their definition writing. Due to the elaborated responses (e.g., “to stop and prevent from something” for the word “thwart”), the experimenter and the observer sometimes accepted different parts of their responses. For accuracy of identification of both parts of speech and synonyms, the IOA was consistently 100%. For accuracy ratings of sentences (generalization), the IOA ranged from 20% to 100% with a
mean of 57.9%. The low agreement occurred exclusively with students 2 (David) and 3 (Rachel), who tended to write illegibly. The experimenter and the observer differed by one point in those cases, in which the experimenter scored 1 point higher than the observer. The difference may be due to the familiarity to the participants and their writing on the side of the experimenter.

*What are the Effects of Distributed Versus Massed Practices Embedded in a Response-Card Activity on the Vocabulary Performances as Measured by Accuracy of Definition Writing?*

During vocabulary probes, students generated their written definition of the five words next to each word written on a probe sheet. A student’s definition writing was considered correct if it matched the original definition sentence or synonyms presented in the instruction slide or if it conveyed the same meaning as the original sentence/synonym presented during instruction and were grammatically correct. It did not have to start with a capital letter in order to be considered correct. Likewise, it could be correct even if the sentence contained grammatical errors or spelling errors as long as the sentence conveyed the original semantic meaning or the spelling errors represented the phonetic characteristics of the original sentence or word. The mean accuracy of definition writing for all students is presented in Figure 4.1, followed by a presentation of the individual student’s results in Figures 4.2 through 4.11.

*Mean Accuracy of Definition Writing*

The mean accuracy of definition writing for the five students is presented in Figure 4-1. The mean score of the five students showed a higher trend in the MP condition on the post-instruction probes (shown in Figure 4.1 with closed circles) in
comparison to the DP condition (shown with closed triangles), especially during the initial four weeks. The mean scores in the DP condition showed a steep ascending trend in the last two weeks (Weeks 6 and 8). On the maintenance probes in the MP condition (shown in Figure 4.1 with open circles), students’ mean scores showed a descending trend whereas the scores remained relatively stable at lower level in the DP condition (shown with open triangles).

Figure 4.1. Mean accuracy of definition writing.

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Student1 (Jessie)

Massed practice (MP) condition. On the post-instruction probes in the MP condition (shown in Figure 4.2 with closed circles), Jessie’s scores on the accuracy of definition writing ranged from 1 to 5 out of 5 points with a mean of 3.0. She scored all available points (5 points) on the first probe but scored 1 out of 5 on week 3’s probe; after the second probe in this condition, her scores showed an ascending trend from 1 out of 5 to 2 out of five in week 5 and 4 out of 5 in week 7. Jessie’s maintenance probe scores for definition writing in this condition (shown with open circles), there were only three data points due to an absence in week 3. Her scores showed a downward trend, with a range of 0 to 3 points and a mean of 1.3.

Distributed practice (DP) condition. On the post-instruction probes for accuracy of definition writing in the DP condition (shown in Figure 4.2 with closed triangles), Jessie’s scores were relatively stable with a range of 2 to 3 out of 5 points and a mean of 2.5. Her scores started with 3 out of 5 points on week 2’s probe; it dropped and stayed at 2 out of 5 points for two consecutive probes (Weeks 4 and 6); finally, it increased back to 3 out of 5 on the probe in week 8. Her maintenance probe scores on accuracy of definition writing in this condition (shown with open triangles) were stable at lower level, with a range of 0 to 1 points and a mean of 0.3. Her scores were at 0 for three consecutive probes (Weeks 2, 4, and 6) before increased to 1 in the last week (Week 8).

Cumulative accuracy. Presentation of the accuracy of definition writing in a cumulative manner, Jessie’s scores for post-instruction probes were slightly higher for the MP condition (shown in Figure 4.3 with closed circles) in comparison to the DP condition (closed triangles). Specifically, in the MP condition, Jessie learned a total of two more
words than in the DP condition. Although the data paths for both conditions were parallel, scores for the MP condition were slightly higher.

*Figure 4.2. Accuracy of definition writing for student 1 (Jessie).

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.3. Cumulative accuracy of definition writing for student 1 (Jessie).

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.

Student 2 (David)

Massed practice (MP) condition. On the post-instruction probes on accuracy of definition writing in the MP condition (shown in Figure 4.4 with closed circles), David’s scores were stable with a range of 0 to 1 out of 5 points with a mean of 0.8. He scored 1 out of 5 points for three consecutive probes (Weeks 1, 3, and 5). In the last probe in this condition (i.e., Week 7), his score dropped to 0 out of 5. His maintenance probe scores on accuracy of definition writing in this condition (shown with open circles) were parallel in an exactly same manner, with a range of 0 to 1 out of 5 points and a mean of 0.8.
Figure 4.4. Accuracy of definition writing for student 2 (David).

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.5. Cumulative accuracy of definition writing for student 2 (David).

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
**Distributed practice (DP) condition.** On the post-instruction probes on accuracy of definition writing in the DP condition (shown in Figure 4.4 with closed triangles), David’s scores were variable but showed a slight ascending trend, with a range of 0 to 2 out of 5 points and a mean of 1.0. His score dropped from 1 out of 5 to 0 out of 5 on the first two probes (i.e., Week 2 and 4); after these two probes, his score increased back to 1 out of 5 in Week 6; in week 8, his score increased further again to 2 out of 5. His maintenance probe scores on accuracy of definition writing in this condition (shown with open triangles) were stable at 0 for all probes.

**Cumulative accuracy.** Presentation of the accuracy of definition writing in a cumulative manner, David’s scores for post-instruction probes for both MP (shown in Figure 4.5 with closed circles) and DP (closed triangles) conditions did not show consistent increase and remained relatively stable. Specifically, in the MP condition, David learned the definitions of a total of 3 words; in the DP condition, he learned the definitions of a total of 4 words.

**Student 3 (Rachel)**

**Massed practice (MP) condition.** On the post-instruction probes in the MP condition (shown in Figure 4.5 with closed circles), Rachel’s scores on accuracy of definition writing ranged from 3 to 5 out of 5 points with a mean of 4.3. She scored 3 out of 5 points on probe in week 1; in the second probe in the MP condition (i.e., Week 3), her score increased to 5 out of 5 points. On the last two probes, Rachel’s scores went up and down; she scored 4 out of 5 points in week 5 and 5 points in week 7. Rachel’s maintenance probe scores on accuracy of definition writing in this condition (shown with open circles) showed a downward trend with a range of 0 to 5 points and a mean of 1.8.
Figure 4.6. Accuracy of definition writing for student 3 (Rachel).

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.7. Cumulative accuracy of definition writing for student 3 (Rachel).

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
Rachel scored 5 points on week 1’s probe, 0 point on week 3’s probe, and 1 point on two consecutive probes (Weeks 5 and 7).

Distributed practice (DP) condition. On the post-instruction probes on accuracy of definition writing in the DP condition (shown in Figure 4.6 with closed triangles), Rachel’s scores ranged from 0 to 4 out of 5 points with a mean of 2.0. She scored 2 out of 5 points on week 2’s probe; on the week 4’s probe, she scored 0 points; after that, her scores increased to 2 points on week 6’s probe and to 4 points on week 8’s probe. Her maintenance probe scores in this condition (shown with open triangles) showed a smaller decreasing trend except for an increase on the last probe with a range of 0 to 2 points and a mean of 1.3. Rachel scored 2 out of 5 points on week 2’s probe; 1 point on week 4’s probe, and 0 point on week 6’s probe. After these three probes, she scored 2 points on the last probe (Week 8).

Cumulative accuracy. Presentation of the accuracy of definition writing in a cumulative manner, Rachel’s scores for post- instruction probes were higher for the MP condition (shown in Figure 4.7 with closed circles) in comparison to the DP condition (closed triangles). Specifically, in the MP condition, Rachel learned a total of 7 more words than in the DP condition. The difference between MP and DP was clear.

Student 4 (Kelly)

Massed practice (MP) condition. On the post-instruction probes on accuracy of definition writing in the MP condition (shown in Figure 4.8 with closed circles), Kelly’s scores ranged from 2 to 5 out of 5 points and a mean of 4.0. She scored all available points (5 points) on two consecutive probes (Weeks 1 and 3); in week 5’s probe in this condition, her score dropped to 2 out of 5, followed by an increase in the score to 4 out of 5 points.
Figure 4.8. Accuracy of definition writing for student 4 (Kelly).

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.9. Cumulative accuracy of definition writing for student 4 (Kelly).

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
5 in week 7’s probe. Her maintenance probe scores on accuracy of definition writing in this condition (shown with open circles) showed a downward trend, with a range of 2 to 5 points and a mean of 3.3. She scored 5 out of 5 points on week 1’s probe, 4 points on week 3’s probe, and 2 points for the remaining two consecutive probes (Weeks 5 and 7).

**Distributed practice (DP) condition.** On the post-instruction probes on accuracy of definition writing in the DP condition (shown in Figure 4.8 with closed triangles), Kelly’s scores ranged from 3 to 5 out of 5 points with a mean of 4.0. Kelly’s score increased for three consecutive probes; she scored 3 out of 5 points on week 2’s probe, 4 points on week 4’s probe, and 5 points on week 6’s probe. After that, Kelly’s score dropped to 4 out of 5 points on week 8’s probe. Her maintenance probe scores on accuracy of definition writing in the DP condition (shown with open triangles) were variable, with a range of 1 to 4 points and a mean of 2.3. Kelly scored 4 points on week 2’s probe; her score dropped to 1 point on week 4’s probe; after these two probes, her scores for the remaining two consecutive probes (Weeks 6 and 8) remained at 2 points.

**Cumulative accuracy.** Presentation of the accuracy of definition writing in a cumulative manner, Kelly’s scores for post-instruction probes in the MP condition (Shown in Figure 4.9 with closed circles) almost exactly paralleled scores in the DP condition (closed triangles). Specifically, Kelly learned a total of 16 words in both MP and DP conditions.

*Student 5 (Sam)*

**Massed practice (MP) condition.** On the post-instruction probes on accuracy of definition writing in the MP condition (shown in Figure 4.10 with closed circles), Sam’s scores were stable with full points (5 out of 5 points) for all four probes. His maintenance
Figure 4.10. Accuracy of definition writing for student 5 (Sam).

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.11. Cumulative accuracy of definition writing for student 5 (Sam).

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
probe scores on accuracy of definition writing in this condition (shown with open circles) showed a downward trend with a range of 2 to 5 points and a mean of 3.8. On the first maintenance probe in this condition (Week 1), he scored 4 out of 5. Then his scores increased to 5 out of 5 on week 3’s probe; after these two probes, his scores decreased to 4 out of 5 on week 5’s probe and to 2 out of 5 points on week 7’s probe.

**Distributed practice (DP) condition.** On the post-instruction probes in the DP condition (shown in Figure 4.10 with closed triangles), Sam’s scores on accuracy of definition writing ranged from 4 to 5 out of 5 points with a mean of 4.3. Sam scored 4 out of 5 points on week 2’s probe, 5 points on week 4’s probe, and 4 points for the remaining two probes (Week 6 and Week 8). His maintenance probe scores on accuracy of definition writing in this condition (shown with open triangles) showed a downward trend, with a range of 2 to 4 and a mean of 2.8. He scored 4 points on week 2’s probe, 3 points on week 4’s probe, and 2 points for two probes (week 6 and 8).

**Cumulative accuracy.** Presentation of the accuracy of definition writing in a cumulative manner, Sam’s scores for post-instruction probes were slightly higher for the MP condition (shown in Figure 4.11 with closed circles) in comparison to the DP condition (closed triangles). Specifically, in MP condition, Sam learned a total of 3 more words than in the DP condition. Although two data paths almost paralleled, the path for the MP condition started to split from the path for the DP condition toward the end.
What are the Effects of Distributed Versus Massed Practices Embedded in a Response-Card Activity on the Vocabulary Performances as Measured by Identification of Parts of Speech?

During vocabulary probes, there were five multiple-choice prompts to which students circled the correct part of speech for each of the five vocabulary words. A response was considered correct if it matched the item in the experimenter-prepared answer key in the appropriate space. The results are organized by student in Figures 4.12 through 4.21.

Student 1 (Jessie)

Massed practice (MP) condition. Jessie’s post-instruction probe scores on accuracy of identification of parts of speech in the MP condition (shown with closed circles in Figure 4.12) showed a downward trend with a range of 0 to 5 and a mean of 2.3. Her score started with 5 out of 5 points on week 1’s probe. However, Jessie’s scores decreased for the remaining three probes. Specifically, she scored 3 out of 5 points on week 3’s probe, 1 point on week 5’s probe, and 0 points on week 7’s probe. The maintenance probe scores on accuracy of identification of parts of speech in this condition (shown with open circles in Figure 4.11) had only three data points due to her absence in week 3. Jessie’s scores showed a downward trend, with 5 points on week 1’s probe and 0 point on weeks 5 and 7’s probes. She was absent on the maintenance probe day in Week 3. Jessie’s maintenance performances had a range of 5 to 0 and a mean of 1.7.

Distributed practice (DP) condition. The post-instruction probe scores on accuracy of identification of parts of speech in the DP condition (shown with closed
Figure 4.12. Accuracy of identification of parts of speech for student 1 (Jessie)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.13. Cumulative accuracy of identification of parts of speech for student 1 (Jessie)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
triangles) were variable at a low level with a range of 0 to 3 and a mean of 1.5. On the first post-instruction probe in this condition (Week 2), she scored 3 points; on week 4’s probe, her score dropped to 0; on week 6’s probe, Jessie scored 1 point out of 5 points, and 2 points on week 8’s probe. The maintenance probe scores on accuracy of identification of parts of speech were stable at low level with a range of 0 to 1 and a mean of 0.3 (open triangles). She scored 0 point on three consecutive probes (Weeks 2, 4, and 6); on week 8’s probe, she scored 1 point out of 5 points.

Cumulative accuracy. Presentation of the accuracy of identification of parts of speech in a cumulative manner, Jessie’s scores on the post-instruction probes were slightly higher for the MP condition (shown with closed circles in Figure 4.13) in comparison to the DP condition (closed triangles). However, the superiority of the MP condition is due to the first probe’s high score and the rate of increase was equally small for both conditions. Specifically, Jessie scored a total of 3 more words in the MP condition than in the DP condition.

Student 2 (David)

Massed practice (MP) condition. David’s post-instruction probe scores on accuracy of identification of parts of speech in the MP condition (shown with closed circles in Figure 4.14) were variable. He scored 5 points on week 1’s probe in this condition. After this, David scored 0 on week 3’s probe; 1 on week 5’s probe, and 0 on week 7’s probe. His performance in this condition ranged from 0 to 5 with a mean of 1.5. The maintenance probe scores on accuracy of identification of parts of speech in this condition (shown with open circles) showed a gradual downward trend. In the maintenance probes in the MP condition, his scores showed a range of 2 to 5 and a mean
Figure 4.14. Accuracy of identification of parts of speech for student 2 (David)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.15. Cumulative accuracy of identification of parts of speech for student 2 (David)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
of 3. He scored 5 points on week 2’s probe. On week 4’s probe, he scored 3 points and 2 points on week 6’s probe. After that, David’s score stayed at 2 points on week 8’s probe.

Distributed practice (DP) condition. The post-instruction probe scores on accuracy of identification of parts of speech in the DP condition (shown with closed triangles in Figure 4.14) were less variable compared to the MP condition with a range of 0 to 2 and a mean of 1.3. He scored 2 points for two consecutive probes (Weeks 2 and 4), followed by decreases to 1 point on week 6’s probe and 0 point on week 8’s probe. The maintenance probe scores on accuracy of identification of parts of speech in the DP condition were variable with a mean of 0 to 5 and a mean of 2 (shown with open triangles). On week 2’s probe, he scored 2 points; on week 4’s probe, his score increased to 5 points; however, on the remaining two probes, his scores showed a downward trend with 1 point on week 6’s probe and 0 point on week 8’s probe.

Cumulative accuracy. Presentation of the accuracy of identification of parts of speech in a cumulative manner, David’s scores on the post-instruction probes in the MP condition (shown in Figure 4.15 with closed circles) were slightly higher in comparison to the DP condition (closed triangles). However, the difference was very small and scores in both conditions increased with a very small increments. Specifically, David learned a total of 1 more word in the MP condition than in the DP condition.

Student 3 (Rachel)

Massed practice (MP) condition. Rachel’s post-instruction probe scores on accuracy of identification of parts of speech in the MP condition (shown with closed circles in Figure 4.16) showed an upward trend with a range of 2 to 4 and a mean of 3. On week 1’s probe, Rachel scored 2 out of 5 points. On week 3’s probe, she scored 3
points, and 4 points on week 5’s probe. On week 7’s probe, Rachel scored 3 out of 5 points. In the maintenance probes in the MP condition, Rachel’s scores on accuracy of identification of parts of speech were more variable with a range of 0 to 5 and a mean of 2.8 (shown with open circles). On week 1’s probe, she scored 5 out of 5 points; on week 3’s probe, her score dropped to 0; on the remaining two probes, her scores increased and remained at 3 out of 5 points (Week 5 and 7).

**Distributed practice (DP) condition.** The post-instruction probe scores on accuracy of identification of parts of speech in the DP condition (shown with closed triangles in Figure 4.16) were low initially and gradually increased with a range of 0 to 2 and a mean of 0.8. She scored 0 points for two consecutive probes (Week 2 and 4) before increasing her scores by 1 point to 1 point on week 6’s probe (Week 6). After that, Rachel scored 2 out of 5 points on week 8’s probe. The maintenance probe scores on accuracy of identification of parts of speech in the DP condition were variable and ranged from 0 to 3 with a mean of 1.5 (shown with open triangles). Rachel scored 0 out of 5 points on week 2’s probe, 2 points on week 4’s probe, and 3 points on week 6’s probe. After these three probes, Rachel’s score dropped to 1 point on week 8’s probe.

**Cumulative accuracy.** Presentation of the accuracy of identification of parts of speech in a cumulative manner, Rachel’s scores on the post-instruction probes were higher in the MP condition (shown in Figure 4.17 with closed circles) in comparison to the DP condition (closed triangles). Specifically, Rachel learned a total of 9 more words
Figure 4.16. Accuracy of identification of parts of speech for student 3 (Rachel)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.17. Cumulative accuracy of identification of parts of speech for student 3 (Rachel)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
in the MP condition than in the DP condition. The difference between the two conditions was clear and widened as weeks continued.

**Student 4 (Kelly)**

*Massed practice (MP) condition.* Kelly’s post-instruction probe scores on accuracy of identification of parts of speech in the MP condition (shown with closed circles in Figure 4.18) were variable at higher level with a range of 3 to 5 and a mean of 4 points. She scored 5 points on week 1’s probe, 3 points on week 3’s probe, 5 points on week 5’s probe, and 3 points on week 7’s probe. In the maintenance probes in the MP condition (shown with open circles), Kelly’s scores on accuracy of identification of parts of speech showed a downward trend with a range of 0 to 5 and a mean of 2.8. Kelly scored 5 points on week 1’s probe, followed by a decrease to 3 points for two consecutive probes (Weeks 3 and 5) and a further decrease to 0 point on week 7’s probe.

*Distributed practice (DP) condition.* The post-instruction probe scores on accuracy of identification of parts of speech in the DP condition (shown with closed triangles in Figure 4.18) were variable with a range of 0 to 5 and a mean of 2.8. She scored 5 out of 5 points on week 2’s probe, 3 points on week 4’s probe, 0 point on week 6’s probe. After these three probes, Kelly scored 3 points on week 8’s probe. The maintenance probe scores in the DP condition showed a range of 2 to 5 and a mean of 3.3 (shown with open triangles in Figure 4.17). She scored 3 points on week 2’s probe, 5 points on week 4’s probe, and 3 points on week 6’s probe; on week 8’s probe, she scored 2 out of 5 points.

*Cumulative accuracy.* Presentation of the accuracy of identification of parts of speech in a cumulative manner, Kelly’s scores on the post-instruction probes split toward
Figure 4.18. Accuracy of identification of parts of speech for student 4 (Kelly)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.19. Cumulative accuracy of identification of parts of speech for student 4 (Kelly)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
the end, showing a higher trend in the MP condition (shown in Figure 4.19 with closed circles) in comparison with the DP condition (closed triangles). Data paths for both condition showed an exact parallel for two probes; from the third probe, Kelly’s score in the MP condition were clearly higher in comparison with the DP condition; specifically, in the MP condition, Kelly learned a total of 5 more words.

**Student 5 (Sam)**

*Massed practice (MP) condition.* Sam’s post-instruction probe scores on accuracy of identification of parts of speech in the MP condition (shown with closed circles in Figure 4.20) were variable with a range of 2 to 5 and a mean of 3.8. He scored 5 points on week 1’s probe, 3 points on week 3’s probe, 5 points on week 5’s probe, and 2 points on week 7’s probe. In the maintenance probes in the MP condition (shown with open circles), Sam’s performance on accuracy of identification of parts of speech almost paralleled his post-instruction probe performances with a range of 3 to 5 and a mean of 4.3. More specifically, Sam scored 5 points on week 1’s probe, 3 points on week 3’s probe, 5 points on week 5’s probe, and 4 points on week 7’s probe.

*Distributed practice (DP) condition.* The post-instruction probe scores on accuracy of identification of parts of speech in the DP condition (shown with closed triangles in Figure 4.20) showed a range of 2 to 5 and a mean of 3.5. Sam scored 3 out of 5 on week 2’s probe; 5 points on week 4’s probe; 2 points on week 6’s probe; and 4 points on week 8’s probe. In the maintenance probes in the MP condition (shown with open triangles), Sam’s performances on accuracy of identification of parts of speech showed a similar pattern with the post-instruction probe scores with a range of 1 to 5 points and a mean of 2.8. Sam scored 1 out of 5 points on week 2’s probe; after that, his
Figure 4.20. Accuracy of identification of parts of speech for student 5 (Sam)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.21. Cumulative accuracy of identification of parts of speech for student 5 (Sam)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
score went up to 5 out of 5 points on week 4’s probe. After this increase, his score dropped to 2 points on week 6’s probe and went up slightly to 3 out of 5 points on week 8’s probe.

**Cumulative accuracy.** Presentation of the accuracy of identification of parts of speech in a cumulative manner, Sam’s scores on the post-instruction probes did not show difference between two conditions, although scores in the MP condition were slightly higher (shown in Figure 4.21 with closed circles) in comparison to the scores in the DP condition (closed triangles). There were 0 to 2 points differences between the data paths for the two conditions. Specifically, Sam learned a total of 1 more word in the MP condition than in the DP condition.

*What are the Effects of Distributed Versus Massed Practices Embedded in a Response-Card Activity on the Vocabulary Performances as Measured by Accuracy of Identification of Synonym?*

During vocabulary probes, there were five multiple-choice prompts to which students circled the letter that corresponded to the correct synonym for each of the five vocabulary words. A multiple-choice option for a word involved a correct option and two distracters. A response was considered correct if it matched the item in the experimenter-prepared answer key in the appropriate space. The mean accuracy of identification of synonyms for all students is presented in Figure 4.22, followed by a presentation of individual students’ results presented Figures 4.23 through 4.33.

**Mean Accuracy of Identification of Synonyms**

The mean accuracy of identification of synonyms for the five students is presented in Figure 4.22. Students’ mean scores were relatively stable with a range of 4.2
to 5.0 out of 5 points and a mean of 4.7. On the post-instructional probes in the MP condition (shown with closed circles in Figure 4.22), students’ mean scores showed an ascending trend and surpassed the initial higher scores in the DP condition (shown with closed triangles) with a range of 4.2 to 5.0 and a mean of 4.6. On the maintenance probes in the MP condition (shown with open circles), students’ mean scores were variable with a range of 4.2 to 5.0 and a mean of 4.8, whereas in the DP condition (shown with open triangles), students’ scores paralleled those in the post-instructional probes in the same condition with a range of 4.2 to 4.9 and a mean of 4.6.

![Figure 4.22. Mean accuracy of identification of synonyms.](image)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Student 1 (Jessie)

**Massed practice (MP) condition.** Jessie’s post-instruction probe scores on accuracy of identification of synonyms in the MP condition (shown with closed circles in Figure 4.23) were stable at high level. She scored 5 points out of 5 for all probes. The maintenance probe scores on accuracy of identification of synonyms in this condition (shown with open circles) also showed a stable and high level of performances with the score of 5 for all probes.

**Distributed practice (DP) condition.** The post-instruction probe scores on accuracy of identification of synonyms in the DP condition (shown with closed triangles in Figure 4.23) were also stable at high level with a range of 4 to 5 and a mean of 4.8. She scored 5 out of 5 points on week 2, 4, and 6’s probes; after these three probes, she scored 4 points on week 8’s probe. The maintenance probe scores on accuracy of identification of synonyms in the DP condition (shown with open triangles) were more variable with a range of 3 to 5 points and a mean of 3.8. Jessie scored 4 out of 5 points on week 2’s probe and 5 points on week 4’s probe; after these probes, her score went down to 3 out of 5 points for two consecutive probes (week 6 and 8).

**Cumulative accuracy.** Presentation of the accuracy of identification of synonyms in a cumulative manner, Jessie’s scores on the post-instruction probes showed a small split toward the end. Specifically, Jessie’s scores paralleled in a exactly same manner until probe 3, but her score in the MP condition (shown with closed circles in Figure 4.24) resulted in a total of 1 more word leaned in comparison to the DP condition (closed triangles).
Figure 4.23. Accuracy of identification of synonyms for student 1 (Jessie)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.24. Cumulative accuracy of identification of synonyms for student 1 (Jessie)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
Student 2 (David)

**Massed practice (MP) condition.** David’s post-instruction probe scores on accuracy of identification of synonyms in the MP condition (shown with closed circles in Figure 4.25) were relatively stable with a range of 4 to 5 and a mean of 4.5 points. He scored 5 points on week 1’s probe, and 4 out of 5 points for two consecutive probes (Week 3 and 5); after these probes, David’s score went up to 5 points on week 7’s probe. The maintenance probe scores on accuracy of identification of synonyms in the MP condition (shown with open circles) were more variable with a range of 3 to 5 and a mean of 4.0 points. He scored 5 points on week 1’s probes and 3 points on week 3 and 5’s probes; after these three probes, his score increased back to 5 out of 5 points on week 7’s probe.

**Distributed practice (DP) condition.** The post-instruction probe scores on accuracy of identification of synonyms in the DP condition (shown with closed triangles in Figure 4.25) showed a downward trend after an initial increase with a range of 3 to 5 and a mean of 3.8. He scored 4 and 5 points on the first two probes (Week 2 and 4); but his scores decreased to 3 out of 5 points on week 6’s probe and remained at the same level on week 8’s probe. The maintenance probe scores on accuracy of identification of synonyms in this condition (shown with open triangles in Figure 4.23) almost paralleled the post-instruction scores with a range of 3 to 5 and a mean of 4 points. In this condition, the last probe (Week 8) yielded a score of 4 points instead of 3 points.

**Cumulative accuracy.** Presentation of the accuracy of identification of synonyms in a cumulative manner, David’s scores on the post-instruction proves were slightly
Figure 4.25. Accuracy of identification of synonyms for student 2 (David)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.26. Cumulative accuracy of identification of synonyms for student 2 (David)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
higher in the MP condition (shown with closed circles in Figure 4.26) in comparison to the scores in the DP condition (closed triangles). David learned a total of 3 more words in the MP condition than in the DP condition.

**Student 3 (Rachel)**

*Massed practice (MP) condition.* Rachel’s post-instruction probe scores on accuracy of identification of synonyms in the MP condition (shown with closed circles in Figure 4.27) showed an upward trend with a range of 1 to 5 and a mean of 3.8. Her scores started with 1 point on week 1’s probe; however, Rachel’s scores increased to 4 out of 5 points on week 3’s probe and increased further more to 5 points on week 5 and 7’s probes.

The maintenance probe scores on accuracy of identification of synonyms in the MP condition (shown with open circles) were less variable with a range of 4 to 5 and a mean of 4.5. Rachel scored 5 points on week 1’s probe; after that, she scored 4 out of 5 points on two consecutive probes (Weeks 3 and 5) and 5 points on week 7’s probe.

*Distributed practice (DP) condition.* The post-instruction probe scores on accuracy of identification of synonyms in the DP condition (shown with closed triangles in Figure 4.27) were stable at the highest level. Rachel scored 5 out of 5 points for all probes. The maintenance probe scores on accuracy of identification of synonyms in the DP condition (shown with open triangles in Figure 4.27) also stayed at high level. Rachel scored 5 out of 5 points for all probes.

*Cumulative accuracy.* Presentation of the accuracy of identification of synonyms in a cumulative manner, Rachel’s post-instruction probe scores in the DP condition (shown with closed triangles in Figure 4.28) were higher in comparison to the MP
Figure 4.27. Accuracy of identification of synonyms for student 3 (Rachel)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.28. Cumulative accuracy of identification of synonyms for student 3 (Rachel)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
condition (closed circles). Specifically, Rachel learned a total of 5 more words in the DP condition than in the MP condition.

*Student 4 (Kelly)*

*Massed practice (MP) condition.* Kelly’s post-instruction probe scores on accuracy of identification of synonyms in the MP condition (shown with closed circles in Figure 4.29) were stable at the highest level. Rachel scored 5 out of 5 points for all probes. The maintenance probe scores on accuracy of identification of synonyms in this condition (shown with open circles) were also stable at highest level.

*Distributed practice (DP) condition.* The post-instruction probe scores on accuracy of identification of synonyms in the DP condition (shown with closed triangles in Figure 4.29) were stable at the highest level (i.e., 5 points) with a small decrease to 4 points on week 7’s probe, resulting in a range of 4 to 5 and a mean of 4.8. The maintenance probe scores on accuracy of identification of synonyms in this condition (shown with open triangles) were again stable at the highest level (i.e., 5 points) for all probes.

*Cumulative accuracy.* Presentation of the accuracy of identification of synonyms in a cumulative manner, Rachel’s scores on the post-instruction probes showed a small split toward the end. Specifically, Rachel’s scores paralleled in a exactly same manner until probe 3, but her score in the MP condition (shown with closed circles in Figure 4.30) resulted in a total of 1 more word leaned in comparison to the DP condition (closed triangles).
Figure 4.29. Accuracy of identification of synonyms for student 4 (Kelly)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.30. Cumulative accuracy of identification of synonyms for student 4 (Kelly)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
**Student 5 (Sam)**

*Massed practice (MP) condition.* Sam’s post-instruction probe scores on accuracy of identification of synonyms in the MP condition (shown with closed circles in Figure 4.31) were stable at the highest level (i.e., 5 points) for all probes. The maintenance probe scores on accuracy of identification of synonyms in this condition (shown with open circles) were also stable at the highest level (i.e., 5 points) for all probes.

*Distributed practice (DP) condition.* The post-instruction probe scores on accuracy of identification of synonyms in the DP condition (shown with closed triangles in Figure 4.31) were stable at the highest level (i.e., 5 points) for all probes. The maintenance probe scores on accuracy of identification of synonyms in this condition (shown with open triangles) were again stable at the highest level (i.e., 5 points) for all probes.

*Cumulative accuracy.* Presentation of the accuracy of identification of synonyms in a cumulative manner, Sam’s performance on the post-instruction probes showed no difference between the two conditions (shown in Figure 4.32).
Figure 4.31. Accuracy of identification of synonyms for student 5 (Sam)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.32. Cumulative accuracy of identification of synonyms for student 5 (Sam)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
What are the Effects of Distributed Versus Massed Practices Embedded in a Response-Card Activity on the Spelling Performances as Measured by the Number of Correctly Spelled Words?

During the vocabulary probes, students wrote the spelling of the words as they heard the experimenter dictate each word. Students’ spelling was scored as correct if it matched the topographical characteristics of the same item in the answer key developed by the first author. A spelling response that did not match the answer key was still counted as correct if it started with a capital letter, contained only capital letters, or contained correct typographical variations (e.g., “a” instead of “ä,” or “g” instead of “ğ”).

A spelling response was counted as incorrect if it contained omission, substitution, insertion, or reversal of any letter(s). The mean accuracy of spelling from all five students is presented in Figure 4.33, followed by presentations of individual results in Figures 4.34 through 4.43.

Mean Accuracy of Spelling

All students’ mean accuracy of spelling is presented in Figure 4.33. After the initial high score, students’ mean scores steadily descended in the MP condition (shown in Figure 4.33 with closed circles). The mean scores in the DP condition (closed triangles) were relatively stable with a range of 3 to 4.8 out of 5 points. On the maintenance probes, students’ scores paralleled those on the post-instructional probes in both conditions (open circles for MP condition, and open triangles for DP condition).

Student 1 (Jessie)

Massed practice (MP) condition. In the post-instruction probes in the MP condition (shown with closed circles in Figure 4.34), Jessie’s accuracy of spelling scores
Figure 4.33. Mean accuracy of spelling.

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
were relatively stable at high level, with a range of 4 to 5 and a mean of 4.8. She scored 5 out of 5 points on week 1’s, followed by a slight decrease to 4 points on week 3’s probe; after these two probes, Jessie scored 5 points on week 5 and 7’s probes. The maintenance probe scores on accuracy of spelling in the MP condition (shown with open circles in Figure 4.34) showed a similar level of performance with the post-instruction probe data with a range of 4 to 5 and a mean of 3.5, although there were only three probe data in this condition due to her absence in week 3. She scored 5 points on week 1 and 5’s probes; on the last maintenance probe in this condition, Jessie scored 4 out of 5 points.

**Distributed practice (DP) condition.** The post-instruction probe scores on accuracy of spelling in the DP condition (shown with closed triangles in Figure 4.34) were stable at the highest level (i.e., 5 points) except for the score on week 8’s probe with a range of 4 to 5 and a mean of 4.8. Jessie scored 5 out of 5 points on week 2, 4, and 6’s probes; after these probes, she scored 4 out of 5 points on week 8’s probe. The maintenance probe scores on accuracy of spelling in the DP condition (shown with open triangles in Figure 4.34) resulted in a descending trend toward the end with a range of 3 to 5 and a mean of 4.3. In the first probe (Week 2), she scored 4 points, followed by 5 points in two consecutive probes (Weeks 4 and 6), and in the last probe, she scored 3 points (Week 8).

**Cumulative accuracy.** Presentation of the accuracy of definition writing in a cumulative manner, Jessie’s performance on the post-instruction probes showed little difference, resulting in the same number of words learned between the two conditions (shown in Figure 4.35).
Figure 4.34. Accuracy of spelling for student 1 (Jessie)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.35. Cumulative accuracy of spelling for student 1 (Jessie)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
**Student 2 (David)**

*Massed practice (MP) condition.* David’s post-instruction probe scores on accuracy of spelling in the MP condition (shown with closed circles in Figure 4.36) showed a downward trend at low level with a range of 0 to 3 and a mean of 1. He scored 3 points on week 1’s probe but decreased to 0 points on two consecutive probes (i.e., Weeks 3 and 5); after these three probes, his score slightly increased to 1 point on week 7’s probe. The maintenance probe scores on accuracy of spelling in the MP condition (shown with open circles in Figure 4.36) showed a similar pattern with a range of 0 to 3 and a mean of 1.3. On week 1’s probe, he scored 3 out of 5 points; after this probe, David’s scores were at lower level, with 1 point on week 3’s probe, 0 point on week 5’s probe, and 1 point on week 7’s probe.

*Distributed practice (DP) condition.* The post-instruction probe scores on accuracy of spelling in the DP condition (shown with closed triangles in Figure 4.36) were at a lower level with a range of 1 to 2 and a mean of 1.8. He scored 2 points on week 2’s probe and 1 point on week 4’s probe; on week 6 and 8’s probe, David scored 2 out of 5 points. The maintenance probe scores on accuracy of spelling in the DP condition (shown with open triangles in Figure 4.36) were more variable but remained at a low level with a range of 0 to 2 and a mean of 1.3. He scored 2 points on week 2’s probe, 0 point on week 4’s probe, 2 points on week 6’s probe, and 1 point on week 8’s probe.

*Cumulative accuracy.* Presentation of the accuracy of definition writing in a cumulative manner, David’s scores for the post-instruction probes were slightly higher for
Figure 4.36. Accuracy of spelling for student 2 (David)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.37. Cumulative accuracy of spelling for student 2 (David)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
the DP condition (shown with closed triangles in Figure 4.37) in comparison with the MP condition (shown with closed circles). More specifically, David learned a total of 4 more words in the DP condition than in the MP condition.

**Student 3 (Rachel)**

*Massed practice (MP) condition.* Rachel’s post-instruction probe scores on accuracy of spelling in the MP condition (shown with closed circles in Figure 4.38) showed a range of 3 to 5 and a mean of 4. She scored 5 out of 5 points on week 1’s probe, 3 points on weeks 3 and 5’s probes; after these probes, Rachel scored 5 points on week 7’s probe. The maintenance probe scores on accuracy of spelling in the MP condition (shown with open circles) showed a similar pattern as the post-instruction probes but at a lower level with a range of 2 to 3 and a mean of 2.5. She scored 3 points on week 1’s probe, 2 points on week 3 and 5’s probes, and 3 points on week 7’s probe.

*Distributed practice (DP) condition.* The post-instruction probe scores on accuracy of spelling in the DP condition (shown with closed triangles in Figure 4.38) were variable with a range of 1 to 4 and a mean of 2.8. Rachel scored 3 points on week 2’s probe but dropped to 1 point on week 4’s probe; after these probes, she scored 4 and 3 points on week 6 and 8’s probes. The maintenance probe scores on accuracy of spelling in the DP condition (shown with open triangles in Figure 4.35) showed a similar pattern as the post-instruction probes but with less variability with a range of 1 and 3 and a mean of 1.8. Rachel’s scores sequenced as 2, 1, 3, and 1 points in each probe.

*Cumulative accuracy.* Presentation of the accuracy of definition writing in a cumulative manner, Rachel’s scores on the post-instruction probes were clearly higher in the MP condition (shown with closed circles in Figure 4.39 in comparison to the DP
Figure 4.38: Accuracy of spelling for student 3 (Rachel)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.39. Cumulative accuracy of spelling for student 3 (Rachel)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
condition (closed triangles). Specifically, Rachel learned a total of 5 more words in the
MP condition than in the DP condition.

*Student 4 (Kelly)*

*Massed practice (MP) condition.* Kelly’s post-instruction probe scores on
accuracy of spelling in the MP condition (shown with closed circles in Figure 4.40)
showed a downward trend with a range of 2 to 5 and a mean of 3.0 points. Kelly scored 5
out of 5 points on week 1’s probe and 3 points on week 3’s probe; on week 5 and 7’s
probes, she scored 2 out of 5 points. The maintenance probe scores on accuracy of
spelling in the MP condition (shown with open circles in Figure 4.40) showed a
descending trend with a range of 1 to 4 and a mean of 2.0 points. Kelly scored 4 points on
week 1) to 2 points in the second probe (Week 3) and 1 points in the third and the last
probes (Weeks 5 and 7).

*Distributed practice (DP) condition.* The post-instruction probe scores on
accuracy of spelling in the DP condition (shown with closed triangles in Figure 4.40)
were variable with a range of 2 to 4 points and a mean of 3.3 points. Kelly scored 4 out of
5 points on week 2’s probe, 3 points on week 4’s probe, and 2 points on week 6’s probe.
After these probes that resulted in a descending trend, her score increased to 4 points on
week 8’s probe. The maintenance probe scores on accuracy of spelling in the DP
condition (shown with open triangles in Figure 4.40) were less variable with a range of 2
to 3 and a mean of 2.5 points. Kelly scored 3 points on week 2’s probe, 2 points on week
4’s probe, 3 points on week 6’s probe, and 2 points on week 8’s probe.

*Cumulative accuracy.* Presentation of the accuracy of definition writing in a
cumulative manner, Kelly’s scores on the post-instruction probes did not show consistent
Figure 4.40. Accuracy of spelling for student 4 (Kelly)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.41. Cumulative accuracy of spelling for student 4 (Kelly)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
difference between the two conditions. Kelly learned 1 more words in the DP condition (shown with closed triangles in Figure 4.41); however, from first probe to the third probe, the MP condition (closed circles) produced a higher trend.

Student 5 (Sam)

*Massed practice (MP) condition.* Sam’s post-instruction probe scores on accuracy of spelling in the MP condition (shown with closed circles in Figure 4.42) were relatively stable with a range of 4 to 5 and a mean of 4.8 points. Sam scored 5 out of 5 points on week 1, 3, and 7’s probes; on week 5’s probe, Sam scored 4 out of 5 points. The maintenance probe scores on accuracy of spelling in the MP condition (shown with open circles in Figure 4.39) showed a stable at the highest level (i.e., 5 points) for all probes.

*Distributed practice (DP) condition.* The post-instruction probe scores on accuracy of spelling in the DP condition (shown with closed triangles in Figure 4.42) were stable at the highest level (i.e., 5 points) for all probes. The maintenance probe scores on accuracy of spelling in the DP condition (shown with open triangles) were relatively stable with a range of 4 to 5 and a mean of 4.8 points. On week 2, 4, and 8’s probes, Sam scored 5 out of 5 points; on week 6’s probe, he scored 4 out of 5 points.

*Cumulative accuracy.* Presentation of the accuracy of definition writing in a cumulative manner, Sam’s scores on the post-instruction probes in DP condition (shown with closed triangles in Figure 4.43) ended with slightly higher (i.e., one more word) data path in comparison to the MP condition (closed circles). However, the two data paths were exactly same for two consecutive probes, and started to split in favor of the MP condition.
Figure 4.42. Accuracy of spelling for student 5 (Sam)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.43. Cumulative accuracy of spelling for student 5 (Sam)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
What are the Effects of Distributed versus Massed Practices Embedded in a Response-Card Activity on the Spelling Performances as Measured by the Percent of Correct Letter Sequences (CLS)?

Students’ spellings of the targeted five words were scored for percent of CLS. Each pair of letters within a word was analyzed whether the sequence was correct (✓) or incorrect (✗) and the number of CLS was divided by the total possible number of correct letter sequences to obtain percent CLS. The results are presented in Figures 4.44 through 4.48.

Student 1 (Jessie)

Jessie’s post-instruction probe scores in both conditions did not show notable difference between the percent CLS, whereas the maintenance scores showed a higher level of performance in the MP condition.

Massed practice (MP) condition. Jessie’s post-instruction probe scores on percent of CLS in the MP condition (shown with closed circles in Figure 4.44) were stable at high level with a range of 91.7 to 100% and a mean of 98.0%. The maintenance probe scores on percent CLS in the MP condition (shown with open circles in Figure 4.41) also showed little variability at high level with a range of 90.0 to 100.0% with a range of 96.7%.

Distributed practice (DP) condition. The post-instruction probe scores on percent CLS in the DP condition (shown with closed triangles in Figure 4.44) were stable at the highest level (i.e., 100%) except for a slight decrease in the last probe (Week 7) to 92.0%. The maintenance probe scores on percent CLS in the DP condition (shown with
Figure 4.44. Percent correct letter sequences (CLS) for student 1 (Jessie)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
open triangles in Figure 4.41) were showed more variability with the range of 64.0 to 100.0% and a mean of 89.5%.

**Student 2 (David)**

David’s percent CLS for post-instruction and maintenance probes for both conditions did not show notable differences. Although a little variability was observed, his percent CLS for all conditions showed approximately similar pattern and means of 50 to 60%.

*Massed practice (MP) condition.* David’s post-instruction probe scores on percent CLS in the MP condition (shown with closed circles in Figure 4.45) had a range of 47.2 to 73.0% and a mean of 56.6%. His percent CLS score was high in the first probe (Week 1) but remained around 50 to 55% for the following three probes (Weeks 3, 5, and 7). The maintenance probe scores on percent CLS in the MP condition (shown with open circles in Figure 4.42) almost paralleled the post-instruction probe scores but had a higher score for the first probe (Week 1). David’s maintenance performances resulted in a range of 43.6 to 95.0% and a mean of 62.2%.

*Distributed practice (DP) condition.* The post-instruction probe scores on percent CLS in the DP condition (shown with closed triangles in Figure 4.42) were relatively stable around 60%. David’s performances on percent CLS in this condition showed a range of 48.6% to 70.0% and a mean of 58.8%. The maintenance probe scores on percent CLS in the DP condition (shown with open triangles in Figure 4.45) paralleled the post-instruction probe scores in this condition except the lower score for the last probe (Week 8). David’s maintenance performances on percent CLS in this condition resulted in a range of 43.2% to 61.1% and a mean of 53.4%.
Figure 4.45. Percent correct letter sequences (CLS) for student 2 (David)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Student 3 (Rachel)

Rachel’s percent CLS for post-instruction and maintenance probes in both conditions showed a better trend for the post-instruction probes than the maintenance probes. Also, the MP condition resulted in a better trend than the DP condition.

**Massed practice (MP) condition.** Rachel’s post-instruction probe scores on percent CLS in the MP condition (shown with closed circles in Figure 4.46) had a range of 87.2% to 100.0%. The maintenance probe scores on percent CLS in the MP condition (shown with open circles in Figure 4.43) showed a lower range of 66.7% to 94.6% and a mean of 83.7%. Her scores decreased from first to the third probe (Week 1 to 5) and backed up on the last probe (Week 7).

**Distributed practice (DP) condition.** The post-instruction probe scores on percent CLS in the DP condition (shown with closed triangles in Figure 4.46) increased toward the end with a range of 62.2% to 89.5% and a mean of 74.4%. The maintenance probe scores on percent CLS in the DP condition (shown with open triangles in Figure 4.43) were stable with a range of 67.6% to 77.8% and a mean of 73.0%.

Student 4 (Kelly)

Kelly’s percent CLS for post-instruction and maintenance probes in both conditions showed higher level of performances for the post-instruction scores compared to the maintenance scores; however, no difference was observed in the level of performances between two conditions.

**Massed practice (MP) condition.** Kelly’s post-instruction probe scores on percent CLS in the MP condition (shown with closed circles in Figure 4.47) had a downward trend and a range of 70.0% to 100.0% and a mean of 82.5%. The maintenance
Figure 4.46. Percent correct letter sequences (CLS) for student 3 (Rachel)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.47. Percent correct letter sequences (CLS) for student 4 (Kelly)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.48. Percent correct letter sequences (CLS) for student 5 (Sam)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
probe scores on percent accuracy of spelling in the MP condition (shown with open circles in Figure 4.47) were relatively stable except for the first probe (Week 1) at lower level with a range of 67.5% to 97.3% and a mean of 76.5%.

**Distributed practice (DP) condition.** The post-instruction probe scores on percent CLS in the DP condition (shown with closed triangles in Figure 4.47) showed a downward trend except for a little increase in the last probe (Week 8). Kelly’s performance in this condition yielded a range of 71.1% to 94.4% and a mean of 83.5%. The maintenance probe scores on percent CLS in the DP condition (shown with open triangles in Figure 4.47) showed a range of 63.0% to 86.1% and a mean of 76.6%. Her performances were relatively stable for three probes (Week 2, 4, and 6) and dropped sharply in the last probe (Week 8).

**Student 5 (Sam)**

Sam’s percent CLS for post-instruction and maintenance probes in both conditions yielded no difference between two conditions; he consistently scored high in all probes except for the third maintenance probe in the DP condition (Week 5), where he scored around 95.0% (shown in Figure 4.48).

What are the Effects of Distributed Versus Massed Practices Embedded in a Response-Card Activity on the Generalization of Vocabulary Knowledge as Measured by the Ratings of Novel Sentences Depicted in a Four-Point Likert Scale?

During the vocabulary probes, students wrote their own sentences using five target words of the week. Each sentence was rated using a 4-point rating scale with 0 being the lowest and 3 being the highest. A rating of 0 was scored if (a) the student’s sentence was unfinished and/or did not contain the target word, (b) there was no response,
or (c) the sentence did not make sense grammatically. A score of 1 was given for sentences that were grammatically correct but too general or if the target word was used in a different context and grammatically incorrect. A score of 2 was given for sentences that were grammatically incorrect but used the target word correctly, or for sentences that were grammatically correct but used the target word in a different context. A score of 3 was given for sentences that used the target word correctly and were grammatically correct. The mean generalization rating scores from all five students are presented in Figure 4.49, followed by individual students’ results in Figures 4.50 through 4.55.

**Mean Ratings of Generalization**

On the post-instructional probes, the mean generalization scores were initially higher in the DP condition (shown with closed triangles in Figure 4.49) in comparison to the scores in the MP condition (closed circles). On the maintenance probes, students’ mean scores in the MP condition were variable with a range of 8.8 to 11.3 (shown with open circles in Figure 4.49); in the DP condition, the mean scores showed a descending trend with a range of 8 to 11.6 (open triangles).

**Student 1 (Jessie)**

*Massed practice (MP) condition.* Jessie’s post-instruction probe scores on accuracy ratings of novel sentences (generalization) in the MP condition (shown with closed circles in Figure 4.50) stayed at lower level with a range of 3 to 6 out of 15 points and a mean of 5.0 points. Jessie scored 3 points on week 1’s probe, 6 points for two consecutive probes (Weeks 3 and 5), and 5 points on week 7’s probe. The maintenance probe scores on accuracy ratings of novel sentence (generalization) in the MP condition (shown with open circles) were a little higher than the post-instruction probe scores with
Figure 4.49. Mean accuracy rating of novel sentences (Generalization).

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
a range of 6 to 8 and a mean of 6.7 out of 15 points, although there were only three probes due to her absence on week 3’s probe. She scored 8.0 points on week 1’s probe; after that, she scored 6.0 points for two consecutive probes (Week 5 and 7).

*Distributed practice (DP) condition.* The post-instruction probe scores on accuracy ratings of novel sentence (generalization) in the DP condition (shown with closed triangles in Figure 4.50) showed an upward trend with a range of 8.0 to 12.0 points and a mean of 10.0 points. Jessie scored 8 out of 15 points on week 2’s probe, 9 points on week 4’s probe, 11 points on week 6’s probe, and 12 points on week 8’s probe.

The maintenance probe scores on accuracy ratings of novel sentence (generalization) in the DP condition (shown with open triangles in Figure 4.46) were variable with a range of 0 to 11.0 points and a mean of 5.3 points. She scored 5.0 points on week 2’s probe and 11.0 points on week 4’s probe; however, her score dropped to 0 point on week 6’s probe, and 5 points on week 8’s probe.

*Cumulative accuracy ratings (Generalization).* Presentation of the accuracy of definition writing in a cumulative manner, Jessie’s performances on the post-instruction probes on accuracy ratings of novel sentence (generalization) in the DP condition (shown with closed triangles in Figure 4.51) were clearly higher in comparison to the MP condition (closed circles). Jessie scored a total of 20 more points in the DP condition than in the MP condition.
Figure 4.50. Accuracy rating of novel sentences (Generalization) for student 1 (Jessie)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.51. Cumulative accuracy rating of novel sentences (Generalization) for student 1 (Jessie)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
Student 2 (David)

Massed practice (MP) condition. David’s post-instruction probe scores on accuracy ratings of novel sentence (generalization) in the MP condition (shown with closed circles in Figure 4.52) were stable with a mean of 7.0 to 8.0 and a mean of 7.3 points. David scored 7 out of 15 points for two consecutive probes (Weeks 1 and 3); after these two probes, he scored 8 points on week 5’s probe, and 7 points on week 7’s probe. The maintenance probe scores on accuracy ratings of novel sentence (generalization) in the MP condition (shown with open circles) showed a slight ascending trend with a range of 6.0 to 8.0 and a mean of 7.8 out of 15 points. David scored 6 out of 15 points on week 1’s probe, 7 points on week 3’s probe, and 8 points on week 5’s probe. On week 7’s probe, David scored 8 points again.

Distributed practice (DP) condition. The post-instruction probe scores on accuracy ratings of novel sentence (generalization) in the DP condition (shown with closed triangles in Figure 4.52) showed a downward trend with a range of 7.0 to 11.0 and a mean of 9.0 out of 15 points. David’s score decreased from 11 points on week 2’s probe, 9 points for two consecutive probes (on Week 4 and 6), and 7 points on week 8’s probe. The maintenance probe scores on accuracy ratings of novel sentence (generalization) in the DP condition (shown with open triangles in Figure 4.48) also showed a downward trend with a range of 3.0 points to 10.0 points and a mean of 7.8 points. David scored 10 out of 15 points for two consecutive probes (Week 2 and 4), 8 points on week 6’s probe, and 3 points on week 8’s probe.

Cumulative accuracy ratings (Generalization). Presentation of the accuracy ratings of novel sentence (generalization) in a cumulative manner, David’s scores on the
Figure 4.52. Accuracy rating of novel sentences (Generalization) for student 2 (David)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.53. Cumulative accuracy rating of novel sentences (Generalization) for student 2 (David)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
post-instruction probes in the DP condition (shown with closed triangles in Figure 4.53) were higher in comparison to the MP condition (closed circles).

**Student 3 (Rachel)**

**Massed practice (MP) condition.** Rachel’s post-instruction probe scores on accuracy ratings of novel sentence (generalization) in the MP condition (shown with closed circles in Figure 4.54) were variable but showed an upward trend with a range of 4 to 13.0 out of 15 points and a mean of 9.8. She scored 4 points on week 1’s probe; after this probe, Rachel’s score increased to 12 points on week 3’s probe, 10 points on week 5’s probe, and 13 points on week 7’s probe. The maintenance probe scores on accuracy ratings of novel sentence (generalization) in the MP condition (shown with open circles) also showed an upward trend with a range of 6.0 to 14.0 points and a mean of 9.8 points. Rachel scored 6 out of 15 points on week 1’s probe. After that, her scores steadily increased. On week 3’s probe, she scored 9 points; on week 5’s probe, she scored 10 points; and on week 7’s probe, she scored 14 out of 15 points.

**Distributed practice (DP) condition.** The post-instruction probe scores on accuracy ratings of novel sentence (generalization) in the DP condition (shown with closed triangles in Figure 4.54) were variable with a range of 6.0 and 13.0 out of 15 points and a mean of 9.3. Her score increased from 12.0 to 13.0 on the first two probes (Wee 2 and 4) but dropped to 6.0 points on week 6’s probe and remained at the same level on week 8’s probe. The maintenance probe scores on accuracy ratings of novel sentence (generalization) in the DP condition (shown with open triangles in Figure 4.54)
Figure 4.54. Accuracy rating of novel sentences (Generalization) for student 3 (Rachel)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.55. Cumulative accuracy rating of novel sentences (Generalization) for student 3 (Rachel)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
Cumulative accuracy ratings (Generalization). Presentation of the accuracy ratings of novel sentence (generalization) in a cumulative manner. Rachel’s scores on the post-instruction probes were initially higher for the DP condition (shown with closed triangles in Figure 4.55) in comparison to the MP practice condition (closed circles) by 5 to 9 points. However, the difference between the two conditions closed toward the end and the MP condition surpassed the DP condition by 2 points.

Student 4 (Kelly)

Massed practice (MP) condition. Kelly’s post-instruction probe scores on accuracy ratings of novel sentence (generalization) in the MP condition (shown with closed circles in Figure 4.56) were variable but showed a downward trend with a range of 11.0 to 15.0 and a mean of 13.0 out of 15 points. She scored 15 points on week 1’s probe, 13 points on next two consecutive probes (i.e., Week 3 and 5), and 11 points on week 7’s probe. The maintenance probe scores on accuracy ratings of novel sentence (generalization) in the MP condition (shown with open circles) were also variable with a range of 11 to 15 points and a mean of 13.3 out of 15 points. She scored 13 points on week 1’s probe and 14 points on week 3’s probe; after these two probes, Kelly’s score dropped to 11 points on week 5’s probe but increased back to 15 points on week 7’s probe.

Distributed practice (DP) condition. The post-instruction probe scores on accuracy ratings of novel sentence (generalization) in the DP condition (shown with
Figure 4.56. Accuracy rating of novel sentences (Generalization) for student 4 (Kelly)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.57. Cumulative accuracy rating of novel sentences (Generalization) for student 4 (Kelly)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
closed triangles in Figure 4.56) were stable for all probes at 12 out of 15 points. The maintenance probe scores on accuracy ratings of novel sentence (generalization) in the DP condition (shown with open triangles) were variable at high level with a range of 11 to 14 and a mean of 12 out of 15 points. Kelly scored 12 points on week 2’s probe, 11 points on week 4’s probe, 14 points on week 6’s probe, and dropped back to 11 points on week 8’s probe.

Cumulative accuracy ratings (Generalization). Presentation of the accuracy ratings of novel sentence (generalization) in a cumulative manner, Kelly’s performances on the post-instruction probes were higher for the MP condition (shown with closed circles in Figure 4.57) in comparison with the DP condition. Specifically, Kelly scored 4 more points in the MP condition than in the DP condition. Although the data paths for the two conditions paralleled, scores for the MP condition were consistently higher.

Student 5 (Sam)

Massed practice (MP) condition. Sam’s post-instruction probe scores on accuracy ratings of novel sentence (generalization) in the MP condition (shown with closed circles in Figure 4.58) were relatively stable at higher level with a range of 13 to 15 out of 15 points and a mean of 14 points. He scored 14 points on week 1’s probe, 15 points on week 3’s probe, dropped to 13 on week 5’s probe, and 14 on week 7’s probe. The maintenance probe scores on accuracy ratings of novel sentence (generalization) in the MP condition (shown with open circles) were variable with a downward trend with a range of 13 to 15 and a range of 14 out of 15 points and a mean of 12.3 points. He scored 14 points on week 1’s probe, 15 points on week 3’s probe dropped to 9 points in the third probe (Week 5), and 11 points in the last probe (Week 7).
Figure 4.58. Accuracy rating of novel sentences (Generalization) for student 5 (Sam)

Note: Closed circles represent post-instruction probe scores from the MP condition, open circles represent maintenance probe scores from the MP condition, closed triangles represent post-instruction probe scores from the DP condition, and the open triangles represent maintenance probe scores from the DP condition.
Figure 4.59. Cumulative accuracy rating of novel sentences (Generalization) for student 5 (Sam)

Note: Closed circles represent post-instruction probe scores from the MP condition and closed triangles represent post-instruction probe scores from the DP condition.
Distributed practice (DP) condition. The post-instruction probe scores on accuracy ratings of novel sentence (generalization) in the DP condition (shown with closed triangles in Figure 4.58) were stable at higher level with a range of 13 to 14 out of 15 points and a mean of 13.5 points. Sam scored 13 points on week 2’s probe, 14 points on week 4’s probe, 13 points on week 6’s probe, and 14 points on week 8’s probe. The maintenance probe scores on accuracy ratings of novel sentence (generalization) in the DP condition (shown with open triangles) were also stable at higher level with a range of 13 to 14 and a mean of 13.8. He scored 14 points for three consecutive probes (i.e., Weeks 2, 4, and 6) and 13 points on week 8’s probe.

Cumulative accuracy ratings (Generalization). Presentation of the accuracy ratings of novel sentence (generalization) in a cumulative manner, Sam’s scores on the post-instruction probes did not show difference between the two conditions (shown in Figure 4.59).

Social Validity

Comparison of Performances Between Students With and Without Disabilities

Comparing the levels of performance in response to massed practice (MP) and distributed practice (DP) for students with disabilities (students 1, 2, and 3) and without disabilities (students 4 and 5), students with disabilities tended to score lower than the students without disabilities. Also, looking at the participants’ performances in a cumulative manner, the rate of increase tended to be greater for students without disabilities compared to those with disabilities. However, student 4’s performances were generally variable and sometimes at a level similar to the students with disabilities.

For the accuracy of spelling, specifically, student 5 (Sam) consistently scored
high on all dependent variables. Student 1 (Jessie) performed at a level similar to student 5, and student 3 (Rachel) was able to perform at a level similar to student 5 in the MP condition. For accuracy of definition writing, only student 3 (Rachel) performed as well as student 5 (Sam) in the MP condition. However, Rachel’s definition writing accuracy in the DP condition was not comparable to student 5 (Sam)’s performance level. Also, other students (students 1 and 2) continued to perform lower than student 5. Identification of parts of speech did not show any difference between students with and without disabilities. Identification of synonyms did not produce differences between students with and without disabilities either, but all students tended to score at a high level; however, student 3’s score was comparable to the other students’ performance levels only in DP condition. For accuracy rating of sentences, performances of students without disabilities (students 4 and 5) were higher than students without disabilities.

**Students’ Ratings on Massed practice (MP) or Distributed practice (DP)**

Nineteen students, including the five target students identified a participants, completed the social validity questionnaire (Appendix F). The results of the students’ ratings are depicted in Table 4.1. To a question that asked about the importance of learning new vocabulary, 13 or 68.4% of students responded that it was “very important,” compared to 4 or 21.1% of students who felt it was “a little” important and 4 or 10.5% of students who responded it was “not at all” important. To a question that asked whether they liked using response cards to learn new words, 9 or 47.4% of the students answered that they did not like it, 6 or 31.6% answered “maybe,” and 4 or 21.1% answered they liked using response cards. To a question that asked the extent of helpfulness of using response cards in learning new words, 6 or 31.6% felt it was not
helpful, 11 or 57.9% answered “maybe,” and 2 or 10.5% answered it was helpful. To questions that asked the helpfulness of either MP or the DP conditions, identical responses were allocated to each rating. That is, 2 or 10.5% of students thought the MP or the DP were not helpful, 8 or 42.1% of students said that the MP or the DP were “a little helpful,” 6 or 31.6% answered that the MP or the DP were “very helpful,” and 3 or 15.8% of students said that there was “no difference.” To a question that asked whether the students liked either the MP or the DP condition, 8 or 42.1% of students said they liked the DP condition, 5 or 26.3% said that they liked the MP, and 6 or 31.6% answered there was “no difference.” To a question that asked which vocabulary-related skill students thought had improved the most, 8 or 42.1% of students chose spelling, 4 or 21.1% chose writing definitions, 4 or 21.1% chose identifying synonyms and parts of speech, and 3 or 15.8% chose writing their own sentences (Table 4.1).

Teacher’s Ratings on Massed practice (MP) or Distributed practice (DP)

The language arts teacher also answered a series of questions about the procedure and the results of the MP and DP vocabulary interventions (see Appendix G). To a question that asked about the ease of preparation for the intervention, she answered it appeared to be easy to implement. When asked about the preparation of the instructional slides and the vocabulary probes separately, she answered they both appeared “very” easy to prepare. When asked about the ease of implementation of the intervention, she responded with a rating scale of 2 with 1 being “not at all” and 4 being “very much.”

To a free-response question about the observed changes in the students’ use of the vocabulary, she noted that the students enjoyed using the newly learned words among
<table>
<thead>
<tr>
<th>Question</th>
<th>Percentage (number) of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Do you think it is important to learn new words?</td>
<td></td>
</tr>
<tr>
<td>No, not at all</td>
<td>10.5 (2)</td>
</tr>
<tr>
<td>A little</td>
<td>21.1 (4)</td>
</tr>
<tr>
<td>Yes, very important</td>
<td>68.4 (13)</td>
</tr>
<tr>
<td>2. Did you like using response cards to learn your words?</td>
<td></td>
</tr>
<tr>
<td>No, not at all</td>
<td>47.4 (9)</td>
</tr>
<tr>
<td>A little</td>
<td>31.6 (6)</td>
</tr>
<tr>
<td>Yes, very much</td>
<td>21.1 (4)</td>
</tr>
<tr>
<td>3. Do you think learning new words using response cards was helpful?</td>
<td></td>
</tr>
<tr>
<td>No, not at all</td>
<td>31.6 (6)</td>
</tr>
<tr>
<td>A little</td>
<td>57.9 (11)</td>
</tr>
<tr>
<td>Yes, very much</td>
<td>10.5 (2)</td>
</tr>
<tr>
<td>4. How helpful do you think presenting the different skills (spelling,</td>
<td></td>
</tr>
<tr>
<td>definition, synonyms, and parts of speech) for one word and then</td>
<td></td>
</tr>
<tr>
<td>moving to the skills for the next word? (DP)</td>
<td></td>
</tr>
<tr>
<td>Not at all</td>
<td>10.5 (2)</td>
</tr>
<tr>
<td>A little</td>
<td>42.1 (8)</td>
</tr>
<tr>
<td>Yes, very much</td>
<td>31.6 (6)</td>
</tr>
<tr>
<td>No difference</td>
<td>15.8 (3)</td>
</tr>
</tbody>
</table>

Table 4.1: Students’ Opinions about MP and DP of Vocabulary (Continued)
Table 4.1 Continued

5. How helpful do you think practicing one of the skills (e.g., spelling) for all five words first and then moving on to the next skill (definition) of the five words? (MP)

- Not at all: 10.5 (2)
- A little: 42.1 (8)
- Yes, very much: 31.6 (6)
- No difference: 15.8 (3)

6. Which did you like better? Practicing different skills about a word (spelling, writing definition, examples or synonyms, and parts of speech), or only one skill for five words and move to the next skill?

   a. learning different things about a word (DP): 42.1 (8)
   b. only one thing first (MP): 26.3 (5)
   c. no difference: 31.6 (6)

7. Which skill do you think has improved the most? Spelling, writing definitions, choosing synonyms and parts of speech, or writing your own sentences?

   a. spelling: 42.1 (8)
   b. writing definition: 21.1 (4)
   c. choosing synonyms and part of speech: 21.1 (4)
   d. writing my own sentences: 15.8 (3)

8. If you were to learn words in the future, which do you want to use?

   a. learning all skills about a word (DP): 47.4 (9)
   b. only one skill for all the words (MP): 15.8 (3)
   c. don’t know: 36.8 (7)
their friends and that they liked to use unusual words such as “banal” and “fetid.” To a question that asked whether she wanted to use the vocabulary intervention in the future, she answered with a rating scale of 3 with 1 being “not at all” and 4 being “very much.” However, when asked about the two intervention conditions (i.e., MP and DP), she did not provide any ratings and indicated that the two did not appear to be different from each other. She also indicated that identification of parts of speech was a hard task for seventh grade students, saying “most students in 7th grade struggle to identify the part of speech because often a noun can be used as an adjective.”
CHAPTER 5

DISCUSSION

The current study attempted to examine the comparative effects of massed-versus distributed practice of vocabulary words embedded in a response card activity on the acquisition, generalization, and maintenance of the vocabulary content skills of seventh grade students with and without disabilities in a general education language arts class. The overall results were mixed for the two instructional conditions (massed practice: MP and distributed practice: DP). In this section, I will provide considerations of the results in terms of general conclusion, followed by discussion of the limitation of the study, points that need to be addressed with repeated investigation in the future, as well as implications for practice for special education and general education teachers.

General Discussion

This study purported to compare the effects of massed versus distributed practices of vocabulary words embedded in a response card practices on acquisition, generalization, and maintenance of vocabulary knowledge of seventh grade students with disabilities (mild mental retardation, speech and language disability, and ADHD) and without disabilities in an inclusive language art classroom. Although some students did show functional relation for selected dependent variables, the results were mixed and
may not lead to a definitive conclusion about the effectiveness of massed and the distributed practice conditions on the participants’ performances. In the following section, implications derived for each dependent variable will be discussed.

Results of the Study

Effects of MP and DP. On the accuracy of definition writing, mean scores of all five students show that the MP condition is slightly better in comparison to the DP condition on both post-instruction and maintenance probes. However, with the ascending trend on post-instruction scores in the DP condition, the results may have shown a different picture if there had been more sessions. Looking at the results by each student, student 1 (Jessie), student 3 (Rachel), and student 5 (Sam) demonstrated a better performances in the MP condition in comparison to the DP condition. These three students were generally enthusiastic during the instructional sessions across practice formats. Therefore, the results may suggest benefit of presenting aspects of vocabulary in the MP format. Student 2 (David) and student 4 (Kelly) showed little difference between the two practice formats on the accuracy of definition writing. They were generally not very enthusiastic during the instructional sessions; therefore, the level of attention may have accounted their performances.

On the accuracy of identification of parts of speech, all students showed higher levels of performances in the MP condition in comparison to the DP condition, although the difference between the two factors were small for student 2 (David) and student 5 (Sam). David was a student who consistently performed lower than other students, and Sam consistently performed better than other students. The language arts teacher indicated this measure as being too difficult for those students; therefore, Sam may have
had prior knowledge of parts of knowledge to show any difference between the two conditions. On the other hand, David seemed to have little prior knowledge of parts of speech and responded randomly on the probes. This measure may have been influenced more strongly by the level of prior knowledge than by the practice formats.

On the accuracy of identification of synonyms, the mean accuracy for all five students did not show a significant difference between the two conditions. Looking at the individual student, performances of student 1 (Jessie), student 2 (David), student 4 (Kelly), and Student 5 (Sam) showed little difference between the two practice formats, although David and Kelly performed slightly higher in the MP condition toward the end of the study. However, I could not tell whether the difference was consistent or by chance due to the limited data points. Student 3 (Rachel)’s performances on the accuracy of identification of synonyms were higher in the DP condition than in the MP condition. Rachel’s performance may suggest a functional relation between the DP and the accuracy of identification of synonyms.

On spelling accuracy, the mean accuracy of spelling for all five students were slightly higher in the DP condition in comparison to the MP condition. Looking at individual student’s performances, little difference was observed for the effects of the two practice formats for student 1 (Jessie), 4 (Kelly), and 5 (Sam). These three students generally performed at higher level on spelling measures. Student 2 (David) showed a higher performance in the DP condition than in the MP condition; however, due to his frequent failure to respond to instruction, the difference may have been by chance. I cannot tell at this point how the DP resulted in higher performance for David. Student 3 (Rachel), on the other hand, performed better in the MP condition on the spelling measure.
She was consistently enthusiastic during the instructional sessions as well as during the probe; that is, she always tried to write all the words on the response cards during instruction and also tried to complete all items on the probe even if she had to work longer. Therefore, I can assume that Rachel’s better performances on the spelling measures in the MP condition were not by chance.

On the accuracy ratings of novel sentences (generalization), the mean accuracy ratings for all students did not show a significant difference between the two conditions, although the scores in the DP condition were slightly higher in comparison to the MP condition. Looking at individual student’s performances, student 1 (Jessie), student 2 (David), and student 3 (Rachel) showed higher performances in the DP condition than in the MP condition. These students have documented IEPs, so despite the varied level of attentiveness during the instructional sessions, their performances may suggest a functional relation between the DP and the generalization measure. For student 4 (Kelly), the MP condition produced slightly higher level of performances than in the DP condition. This may suggest a functional relation between the MP and the generalization measure, but the difference between the conditions may not be significant enough given the limited number of data points available. For student 5 (Sam), there was no difference between the two practice formats on this measure.

Social validity measures. Students’ responses to the questionnaire generally showed no difference between the two practice formats. The same number of students indicated that both MP and the DP formats were helpful; however, 42% of the students preferred the DP format. Although the two practice formats may not influence students’ performances, DP may influence how students perceive vocabulary instruction. With
regard to the use of response cards, about half (i.e., 47.4%) of students indicated that they did not like using the response cards and 31.6% of students indicated that response cards were not helpful in learning vocabulary. School events such as other tests and lack of connection of the vocabulary words to the curriculum activities may have influenced the students’ perception of the vocabulary instruction in conjunction with the use of response cards.

Limitations

The current study has several limitations. In the following section, limitations associated with the procedure of the study and data collection will be discussed.

Procedural Limitations

Comparability of sessions. This study was conducted for a limited number of sessions when many school events took place simultaneously. For this reason, throughout the duration of the study, several factors were present that endangered the comparability of the two conditions. Those factors included (a) presence and/or movement of personnel, (b) school events, (c) timing of maintenance probe administration, and (d) seating arrangements.

First, a student teacher was present in the classroom in addition to the language arts teacher during weeks 1 through 3. Although the impact of the presence of the student teacher is unknown, he did talk to some students when instruction for this study was taking place in order to collect students’ assignments. On some occasions, his movement during the instruction may have distracted the students’ attention.

Second, the intervention coincided with testing that the school had scheduled for the students. For example, in week 3, the students had a unit vocabulary quiz
independently prepared by the language arts teacher. Because of the strong contingency associated with performance on the teacher-prepared vocabulary quiz (i.e., the grade), students tended to focus more on studying for this quiz instead of the vocabulary probe. In addition, the last four weeks (i.e., weeks 5 through 8) coincided with practice tests in preparation for the statewide achievement tests. Therefore, students took another test after each of the vocabulary probes. All of these tests could have confounded with the students’ performances on the vocabulary probes for this study. For example, in week 3, student 4 (Kelly) remarked, “Oh no, we have two vocabulary quizzes today,” when she entered the language arts classroom and saw the envelope that contained the vocabulary probe on her desk.

Third, the timing of the maintenance probes differed in some weeks, resulting in assessment of different lengths of maintenance (i.e., three-day versus two-day maintenance). In weeks 2 and 6, the maintenance probe happened two days after the instruction because of the no-school days due to snow or spring break.

Fourth, the seating arrangements changed frequently during the study, resulting in varied contingencies for students, especially for student 2 (David). For example, student 2 sat closer to girls during the instructional session in week 1 and paid close attention to the instruction. In week 4, seating arrangement changed and he started to sit behind his best friend. After that arrangement was made, he started to chat with his friend during the instruction and seldom responded to the instruction unless the experimenter stood close to him. Also during the maintenance probe in week 7, all students moved from their regular classroom to a computer room to finish their essay writing, which may have affected some of the students’ performance on the vocabulary probe because the
seating arrangement enabled them to look at peers’ responses if they wanted. Although the experimenter monitored the students to prevent them from “cheating” during the probe, there was no guarantee that she prevented all instances of this. In the current study, although maximum efforts were made to keep all the factors for each condition identical, experimentally controlling the instances of the other testing and the school events was extremely difficult. Therefore, conditions may not have been comparable to each other.

Selection of target words. In the current study, no pretest was administered to assess the students’ level of prior knowledge on the target words. The target words were selected from SAT and ACT vocabulary words (in consultation with the language arts teacher), excluding the words that have been taught previously during the school year. However, it is possible that the students already knew some of the words or had prior exposure to them. Similarly, there were no measures available to restrict students’ exposure to the target words outside the classroom except for asking the language arts teacher not to use these words in other occasions. Although the language arts teacher anecdotally reported that she did not use the target words during instruction, students may have had additional incidental exposures to those words through different media. In addition, in week 1, targeted words involved only adjectives. During the instruction, one of the students noticed the pattern and mentioned it to the whole class (i.e., said, “All those words are adjectives”). Therefore, students earned high scores on identification of parts of speech task for week 1.

Another issue associated with the selection of target words was that the words were irrelevant to the on-going instruction in the language arts classroom. The language arts teacher usually selected vocabulary words from the passages she was teaching. In
this study, however, the target words were selected from the SAT or OAT words that had little connection to the activities took place in the language arts class. Therefore, the contingencies for students to learn these words may have been weakened especially when there were other tests simultaneously taking place as the study. These factors may account for the low social validity associated with the use of response cards in this study. Given that the response cards were used consistently in both conditions, they served as a discriminative stimulus for some extra work.

*Design of the vocabulary probes.* The current study involved acquisition measures of the vocabulary knowledge (i.e., definition, identification of part of speech, and identification of synonyms), acquisition measures of spelling (accuracy of spelling and percent CLS), generalization, and maintenance measures. Although there were a total of 35 points per probe, each dependent variable was allocated 5 possible points. Therefore, students’ performances could easily reach the ceiling for each aspect or the dependent variable.

In the current study, due to the limited time allowed for the study, long-term maintenance measures were not available. Given that vocabulary knowledge becomes truly functional when one maintains the knowledge for an extended period of time (Silverman, 2007), long-term maintenance measures may have strengthened the results of the study. In order to take the long-term generalization data, however, the amount of the students’ exposure to the target words should be more strictly controlled. The generalization measure may also warrant some considerations. In much of the existing literature on vocabulary instruction, experts imply a strong relationship between students’ vocabulary knowledge and their reading and/or comprehension (NRP, 2000). In the
current study, generalization was measured by having students write their own sentences using the target words (i.e., expressive written task). Although this measure did suggest how well the students retained the target words, it may not be a sufficient measure to suggest how well students could read and comprehend a novel passage that contain the target words.

**Experimental design.** In the current study, an alternating treatment design was used due to limited schedule. Due to the rapid alternation of conditions characteristic to this design, comparison of the two intervention conditions may have become even more difficult. In addition, there was only one day of instruction for each condition. There should have been more exposure to the week’s vocabulary words.

**Format of instructional sessions.** Related to the amount of exposure to the week’s vocabulary words, there is an issue of whether students should have been taught to mastery of the target words or have been granted equal amount of instruction. In this study, students received the instruction once and their retention of these words was probed. Given the complex nature of vocabulary knowledge that relies heavily on contextual information, this instructional format may have given highly condensed vocabulary information, which may have not necessarily fostered true understanding of the target words among students. In other studies that taught vocabulary to students with mild disabilities, however, target words were taught to mastery. For example, Ross and Stevens (2003) taught spelling of social studies vocabulary words to three elementary students with mild disabilities using 5 second constant time delay within a multiple-baseline across word set context. In this study, three different sets of words were taught until students were able to spell all the words for 100% accuracy for at least two
consecutive training sessions, followed by probes of these words. Comparing the speed to reach mastery in the massed or distributed practice formats may have provided different information for the effectiveness of these two practices.

In addition to the amount of exposure to the target words, this study did not allow for counterbalancing of the words assigned for each condition. Given that the instructional sessions were implemented in a whole class format, all students were exposed to the same words. Although the experimenter made sure that the words taught each week contained similar number of letters, they were not analyzed in terms of difficulty level associated with phonemic rules or spelling rules. Therefore, the words assigned for each practice formats may not have been equal for phonemic difficulty level or for regularity of the spelling. It is possible that the words assigned to the MP condition, for example, were easier than the words assigned to the DP condition in terms of these factors and therefore may have affected the students’ performances in this condition positively.

**Limitations Associated with Data Collection**

_Procedural integrity._ Due to the schedule of the second experimenter, only two of the instructional sessions and probe sessions received procedural integrity check. Throughout the course of the study, many events took place that may have interfered with the comparability of the two conditions; therefore, it was imperative that the intervention and the probe were carried out with integrity. Although the experimenter made sure that she carried out all the steps by referring to the procedural checklists (Appendixes H and I), more frequent scoring of the procedural integrity by an independent scorer may have strengthened the believability of the experimenter’s report.
Students’ response to instruction. The current study did not conduct an observation of students’ behavior during the instruction to measure the level and accuracy of their response card responses. However, anecdotal data of student 2 (David) suggested inconsistent level of responses to the instruction in some sessions or different levels of accuracy of the response card responses. Data on how they responded to each instructional antecedent or how many practices students required until they all responded correctly during the response card practice may enhance the validity of the data.

Relation to Existing Research

Acquisition measures

For accuracy of spelling, there was no difference between the two conditions of vocabulary intervention: massed (MP) and distributed practices (DP). This general pattern was similar for both accuracy of spelling and percent correct letter sequences (CLS). However, Student 1 (Jessie), a student with mild mental retardation, scored as well as Student 4 (Kelly) and Student 5 (Sam) in both conditions. For accuracy of definition writing, there was a general superiority of the MP over the DP condition. The superiority of the MP was also observed for the accuracy of identification of parts of speech and identification of synonyms. However, Student 3 (Rachel) scored higher in the DP only with the identification of synonyms. Factors that contributed to Rachel’s higher performance in the DP condition remain unknown. As for identification of parts of speech, this task itself may be too difficult for the participants. There should have been some additional strategy instruction of how to discriminate and identify parts of speech before the end of the study.
Maintenance measures

The overall superiority of MP over DP parallels with the existing research in which acquisition was enhanced by systematic massed practice (Carnine, Silbert, & Kame’enui, 1990). However, for maintenance, superiority of the DP generally observed in the previous studies was not replicated. In the current study, MP generally produced better performances; also in some weeks, maintenance scores were better than the post-instruction probe scores. This may be due to the repeated exposure to the probe items through which students had an opportunity for observational learning (Ross & Stevens, 2003). Given that students were not taught to mastery levels during the instruction, they could have been considered as still being in the acquisition stage when the instruction ended. Therefore, taking the post-instruction probe could have served as additional opportunity to “learn” the aspects of vocabulary.

Generalization measure

For generalization, students with disabilities (Students 1 through 3) showed a consistent superiority of the DP condition, while students without disabilities (Student 4: Kelly, Student 5: Sam) did not show a consistent pattern. Kelly’s scores were slightly higher on the generalization measure in the MP condition, and Sam performed equally well on the generalization measure in the DP condition. Although this appears consistent with the research on massed or distributed practice and design of DI programs (Carnine, Silbert, & Kame’enui, 1990), generalization measures used in this study (i.e., writing novel sentences using the target words) were not directly targeted during instruction. With a minimal level of instruction (i.e., verbally read an example sentence once), participants were able to produce their own sentences. Another consideration is the
quality of participants’ sentences. In this study, students’ sentences were rated using a four-point Likert scale of 0 to 3; therefore, a score of 3 indicated highest quality for a sentence. Throughout the course of the study, the frequency of full points for each student did not change. This may reflect the possibility of the students having reached the ceiling for the generalization measure.

Curriculum-based measure

The percent correct letter sequences (CLS) used in this study posed another area of consideration. In this study, percent CLS was employed in an attempt to demonstrate the utility of an adapted curriculum-based measurement system that can detect small improvements of students’ performances in comparison to the frequency of accurate responding. However, in this study, students’ spelling performances measured using the frequency counts (i.e., number of correct spelling) and the percent CLS paralleled; percent CLS did not necessarily provide a unique picture of the students’ spelling performances. One possible implication of the use of percent CLS as an indicator of students’ spelling performances may be with lower-performing students. For example, David, who performed consistently lower than other students, scored 3 out of 5 points on both the post-instructional probe and maintenance probe in week 1. However, the percent CLS indicated that David’s CLS on week 1’s maintenance probe (shown with open circles in Figure 4-42) was significantly higher (i.e., 95%) than on the post-instructional probe (i.e., 73%; closed circles in Figure 4-42). In this case, percent CLS may have been more sensitive to small differences to David’s performance as Hosp and Hosp (2003) suggested. However, for students who generally perform well (e.g., Kelly and Sam), the percent CLS may not be necessary.
Social validity

In the previous studies that used response cards, students generally preferred using response cards (e.g., Clayton & Woodard, 2007). In those studies, response cards were used as the primary independent variable, which was compared with a traditional instructional strategy such as hand-raising or in-class discussion. In this study, however, response cards were consistently implemented as part of the intervention across two practice formats. Because of the lack of connection of the vocabulary words to the curriculum activities and the school events simultaneously took place during the study, I can suspect that the intervention was becoming aversive to the students. These factors may account for why response cards received low social validity ratings from students in this study.

The teacher’s ratings to support the ease of preparation and implementation of the vocabulary instruction in this study, however, may support findings from previous studies. Heward (1994) wrote that the “low-tech” ASR strategies such as response cards could easily conform to any given instructional arrangements. The language arts teacher’s ratings in this study suggested that response cards could be one of the strategies that have high level of usability as well as effectiveness.

Implications for Future Research

Research on vocabulary instruction is lacking in an operational definition of vocabulary knowledge, consistent measures of vocabulary performance, and ways to provide effective and efficient instruction. This study is a preliminary attempt to address these shortcomings of vocabulary instruction by incorporating (a) a response card strategy as a way to increase active student response (ASR) and (b) vocabulary skills
presented in either massed or distributed formats. The ultimate purpose of this study was to identify an effective and efficient instructional format for older students.

Systematic replication of this study should attempt a more strict control of the instructional conditions. More specifically, teachers present in the room, seating arrangements, and maintenance timing should be consistent so that conditions are strictly comparable. Systematic replication should also address procedural issues. First, participants’ knowledge on the target vocabulary words should be assessed by implementing a pretest. Second, target words should be equal not only to the number of letters, but also to the phonemic and spelling rules so the difficulty level of those words will be as comparable as possible. Third, there should be a stronger connection between the target words and the on-going instruction. In their analysis of spelling instruction for three elementary students, Ross and Stevens (2003) taught spellings of vocabulary words prior to using them in an inclusive social studies class. The same consideration should be given so that there is a stronger contingency for students for learning the words and that there is less probability for students to guess their answers when they take the vocabulary probes. Fourth, vocabulary probe should be reconstructed so that the number of items for each dependent variable does not easily reach a ceiling. Fifth, long-term maintenance measures of the knowledge of the vocabulary words should be included, in which half of the items come from the MP condition and the remaining half comes from the DP condition. Sixth, an additional generalization measure should be included in which students would receptively identify the appropriate definition or meaning in a novel written passage. Lastly, students’ responses during the instructional session should be recorded so that the number of active responses during instruction can be identical.
Another consideration is related to the experimental design and instructional format. First, in order to ensure enough exposure to each instructional condition and to examine the utility of the response card strategy in vocabulary instruction, an alternating treatment design nested within a reversal design should be used so that the baseline phase of the reversal would be no-response card condition. The massed and distributed practice formats should be alternated in the intervention phases. Instructional sessions should also ensure enough exposure to target vocabulary words. Therefore, there should be multiple instructional sessions so students can reach a pre-determined mastery criterion as illustrated in Ross and Stevens (2003). Specifically, instruction should be done for at minimum of 3 days. In the MP condition, one of the aspects should be taught for all words, such as spelling of five words in a day followed by definition writing of all words in the second day. In the DP condition, aspects for each word should be taught in a day, such as spelling for one of the word, followed by definition writing, identification of part of speech and synonym of the same word in a day. Also, instruction should be given to two different groups while counterbalancing the order in which each group is exposed to the two instructional formats (i.e., MP or DP); specifically, group A would be taught the target words in the MP condition, while group B learn the same words in the DP condition. By doing so, we can minimize the risk of words assigned to one of the practice formats being easier than the words in other format.

Also, given the complex nature of vocabulary knowledge, the construct of vocabulary probes awaits further validation. Although the dependent variables used in this study did cover many aspects of vocabulary knowledge, there remains a question of what role these aspects play in the whole picture of development of vocabulary
knowledge and, ultimately, reading comprehension and written expression. Therefore, research should identify dependent variables that most reliably correlate with students’ true understanding of words and their skill in applying this understanding.

**Implications for Practice**

The instruction used in this study was short and simple. The instruction took 10 to 20 minutes each day, and preparation of the instructional materials was easy. Given the complex nature of vocabulary knowledge (Silverman, 2007), the ease of implementation of the short intervention has strong implications for practice, especially when the intervention allowed for instruction on all different aspects of vocabulary knowledge. Available studies on vocabulary instruction lack a consistent definition of “vocabulary knowledge” and hence lack a consensus as to how vocabulary knowledge is measured. This study identified multiple aspects of vocabulary knowledge, provided time-efficient instruction of all aspects, and demonstrated how these aspects could be measured in a manner that is easy to implement and score (i.e., paper-and-pencil tests). The language arts teacher’s ratings on the ease of preparation and implementation of the instructional slides add a strong practical value to this intervention. Another notable point is the generalization measure that was collected after very little instruction. Students were able to use the words in their novel sentences, and language arts teacher’s anecdotal data suggested that they continued to use the words in their daily communication.

This study suggested another area in which response cards can be used efficiently at middle school level. Response cards have been used with elementary students in seven studies (Armendariz & Umbreit, 1999; Berrong et al., 2007; Christle & Schuster, 2003; Gardner et al., 1994; Godfrey et al., 2003; Maheady et al., 2002; Narayan
et al., 1990) and with college students in four studies (Clayton & Woodard, 2007; Kellum et al., 2001; Marmolejo et al., 2004; Shabani & Carr, 2004). Three previous studies have used response cards with secondary students (Cavanough et al., 1996; Davis & O'Neill, 2004; Horn et al., 2006). Among these three studies, no studies used response cards for vocabulary activities. Therefore, this study posed a new application of response cards to vocabulary acquisition of middle school students with and without disabilities in a general education setting. Although recommendations for vocabulary instruction contented that vocabulary should be taught directly and indirectly using different instructional formats (NRP, 2000), complexity of vocabulary knowledge does not always allow teachers to implement systematic and extensive vocabulary instruction (Jenkins & Dixon, 1983). If teachers use response cards, direct instruction of vocabulary across different aspects can be done efficiently because response cards enable teachers to see the performances of each student in the group. Massed and distributed practices used in this study can also enhance the ease of vocabulary instruction in classroom. As the results of this study indicated, there may be little impact of how the vocabulary aspects are presented on students’ learning of vocabulary. Therefore, teachers could simply teach the aspects of vocabulary in any manner they choose to do it.

*Implications for general education teachers*

Middle school students learn an extensive amount of vocabulary words in many subject areas (Anderson, Wilson, & Fielding, 1988). General education teachers need not only to teach the content but also to teach vocabulary so students can succeed in their classes. Vocabulary instruction becomes even more important if students with disabilities are included in general education classes. The intervention package used in this study was
easy to prepare and implement to the whole class. If teachers can ensure sufficient connection between the curriculum and the vocabulary taught, the intervention package can be a good practice in general education classrooms.

**Implication for special education teachers**

An advantage of the intervention package used in this study is that it could be implemented in large-group or small-group contexts. Also, the vocabulary probes were quick to implement. Special education teachers can use the intervention package in a resource room to provide supplemental preview-type instruction of vocabulary words that are to be covered in the general education classes, regardless of the subject areas, because of the characteristics of response cards; response cards do not limit the subject areas in which they are applied. For example, this intervention package can be used to teach science or social studies vocabulary words.

**Summary**

Extensive vocabulary knowledge is critical to a student’s academic and social performance in school. However, vocabulary is a complex skill that overlaps with many different skills and can be conceptualized with different dimensions. While there is a pressing need to provide effective and efficient vocabulary instruction at all levels, the perceived complexity of vocabulary creates confusion about how to operationalize vocabulary knowledge, how to consistently and systematically measure vocabulary improvements, and how to teach different dimensions of vocabulary. Available research suggests that students should (a) be actively engaged in instruction and (b) be allowed repeated practice opportunities. The current study examined the comparative effects of massed practice (MP) and distributed practice (DP) of aspects of vocabulary (i.e.,
spelling, writing definitions, identifying parts of speech, identifying synonyms, sentence writing) embedded in a response card activity on the vocabulary and spelling acquisition, maintenance, and generalization of five seventh grade students with and without disabilities enrolled in an inclusive language art classroom. The results were mixed, so definitive trends could not be determined. Despite the study’s limitations, there were several implications for future research and suggestions for classroom practice identified.
REFERENCES


Heward, W. L. (2005). Reasons applied behavior analysis is good for education and why those reasons have been insufficient. In W. L. Heward, T. E. Heron, N. A. Neef et al. (ed.). *Focus on Behavior Analysis in Education: Achievement, Challenges, and Opportunities*. Upper Saddle River, NJ: Parentice Hall/Merrill.


195


APPENDICES
APPENDIX A

RESEARCH PARTICIPATION CONSENT FORM
Participation in Educational Research

Provided that prior consent has been provided by the parents, I will allow the students in my [grade] students to be provided with supplemental vocabulary instruction. I understand that a researcher from The Ohio State University will be working with the students for 5-10 minutes, four days a week.

I realize this study will be conducted over a period of 14 to 16 weeks by Moira Konrad, Ph.D. and Madoka Itoi, an OSU Ph.D. student. The purpose of this research is to examine the comparative effects of distributed versus massed vocabulary practices embedded in a response card activity.

I understand that the children’s identity will not be revealed in any publication, document, or any other form of report developed from this project. Additionally, I understand that I may withdraw consent for my students’ participation at any time without penalty.

__________________________
Classroom teacher

__________________________
Principal
APPENDIX B

PROBE SHEET FOR DEFINITION WRITING
Sheet 2

Direction: Write the meaning of each word.

1. adversary means _______________________________________

2. rail means ______________________________________________

3. allude means _____________________________________________

4. erudite means _____________________________________________

5. lofty means _______________________________________________
APPENDIX C

PROBE SHEET FOR PARTS OF SPEECH AND SYNONYM IDENTIFICATION
Sheet 3
Direction: Circle the appropriate ones for each word.

1. allude is a/an (verb, noun, adj.).
   Another word for allude is
   [ a. supply b. imply c. show ]

2. rail is a/an (verb, noun, adj.).
   Another word for rail is
   [ a. criticize b. praise c. run ]

3. lofty is a/an (verb, noun, adj.).
   Another word for lofty is
   [ a. wide b. raise c. towering ]

4. adversary is a/an (verb, noun, adj.).
   Another word for adversary is
   [ a. enemy b. elite c. advice ]

5. erudite is a/an (verb, noun, adj.).
   Another word for erudite is
   [ a. wise b. wary c. wish ]
APPENDIX D

PROBE SHEET FOR SPELLING
Direction: Write the spelling of the word when you hear the teacher says it.

1. _______________________________

2. _______________________________

3. _______________________________

4. _______________________________

5. _______________________________
APPENDIX E

PROBE SHEET FOR GENERALIZATION (SENTENCE WRITING)
Sheet 4

Direction: Using each word, write your own sentence.

1. erudite

_______________________________________________________

_______________________________________________________

2. allude

_______________________________________________________

_______________________________________________________

3. rail

_______________________________________________________

_______________________________________________________

4. adversary

_______________________________________________________

_______________________________________________________

5. lofty

_______________________________________________________

_______________________________________________________
Please circle your answer to each question.

1. Do you think it is important to learn new words?

1  2  3
No, not at all  A little  Yes, very important

2. Did you like using response cards to learn your words?

1  2  3
No, not at all  Maybe  Yes, very much

3. Do you think learning words using response cards was helpful?

1  2  3
No, not at all  Maybe  Yes, very helpful

Now, I will ask about how you practiced the five words.

4. How helpful do you think presenting the different skills (spelling, definition, synonyms, and parts of speech) for one word and then moving to the skills for the next word?

1  2  3  4
No, not at all  A little  Yes, very much  No difference

5. How helpful do you think practicing one of the skills (e.g., spelling) for all five words first and then moving on to the next skill (definition) of the five words?

1  2  3  4
No, not at all  A little  Yes, very much  No difference

6. Which did you like better? Practicing different skills about a word (spelling, writing definition, examples or synonyms, and parts of speech), or only one skill for five words and move to the next skill?

   a. learning different things about a word
   b. only one thing first
   c. no difference

7. Which skill do you think has improved the most? Spelling, writing definitions, choosing synonyms and parts of speech, or writing your own sentences?
8. If you were to learn words in the future, which do you want to use?

   a. learning all skills about a word
   b. only one skill for all the words
   c. don’t know

   a. spelling
   b. writing definition
   c. choosing synonyms and part of speech
   d. writing my own sentences
APPENDIX G

TEACHER QUESTIONNAIRE
Teacher Questionnaire

Please rate the following questions with 1 being the lowest (not at all) and 4 being the highest (very much).

1. Based on your observation of the vocabulary intervention, did it appear to be easy to prepare?
   
   1  2  3  4
   
   a. Please rate the perceived ease of preparing the presentation slides.
      
      1  2  3  4
   
   b. Please rate the perceived ease of preparing the vocabulary tests.
      
      1  2  3  4

2. Based on your observation of the vocabulary intervention, did it appear to be easy to implement?
   
   1  2  3  4

3. Did you notice any changes in the use of vocabulary words in your students? Please give examples.

   ____________________________________________________________
   ____________________________________________________________
   ____________________________________________________________

4. Would you like to use the vocabulary instruction package in your class in the future?
   
   1  2  3  4
   
   a. Would you like to use the distributed skill practice format (i.e., practicing different skills such as spelling, definition, or synonym about a word and move on to the next word) to teach vocabulary?
      
      1  2  3  4
   
   b. Would you like to use the massed skill practice format (i.e., practicing one of the skills such as spelling for all target words first and then move on to the next skill such as definition of all words) to teach vocabulary?
      
      1  2  3  4

5. Please include any comments below.

   ____________________________________________________________
   ____________________________________________________________
   ____________________________________________________________

   216
APPENDIX H

PROCEDURAL INTEGRITY (INSTRUCTION) CHECKLIST
Procedural Integrity (Instruction) Checklist

____ Experimenter distributed a write-on response cards and a marker to the students
____ Experimenter had the right format (distributed or massed) PowerPoint presentation ready on the screen

*Model slides*

1. _____ Experimenter read the word once.
2. _____ Experimenter had the students say the word aloud.
3. _____ Experimenter read the information on the slide once.
4. _____ Experimenter had the students read the sentence aloud once.
5. _____ Experimenter read the example sentence.
6. _____ Experimenter prompts the students to repeat the definition.
7. _____ Experimenter had the students say the part of the speech
8. _____ Experimenter prompted students to spell the word on the response card as saying the spelling out loud.

*Spelling prompt*

9. _____ Experimenter inserted prompts for students to write the word on the response cards

10. _____ Experimenter gave appropriate signal “Boards up!”

11. _____ Experimenter gave the correct consequence

   ① Praise “Good job, the answer is __________”

   ② Corrective feedback if one of the target students responded incorrectly: “The answer is __________. Correct your spelling now.”
**Definition Prompt**

12. ______ Experimenter prompted students to write the definition on the response cards

13. ______ Experimenter gave appropriate signal “Boards up!”

14. ______ Experimenter gave the correct consequence

   ① Praise:

   ② Corrective feedback if one of the target students responded incorrectly: “The answer is ______.”

**Synonym and Part of Speech Prompt**

15. ______ Experimenter prompts the students to write the appropriate letter and an alphabet on the response cards with the appropriate signal, “Boards up”

   ① Praise:

   ② Corrective feedback if one of the target students responded incorrectly: “The answer is ______.”

______ Experimenter went through all the slides

______ Experimenter thanked the students for participating

______ Experimenter collected the response cards and markers

/20
APPENDIX I

PROCEDURAL INTEGRITY (PROVE) CHECKLIST
Procedural Integrity for the Probe

_____ Experimenter prompted the students to take out a pencil and an eraser.

_____ Experimenter prompted the students to take out the first sheet from the envelop

_____ Experimenter dictates five words every 10 seconds

1. [word 1]
2. [word 2]
3. [word 3]
4. [word 4]
5. [word 5]

_____ Experimenter prompted the students to put the first sheet back in the envelop

_____ Experimenter prompted the students to take out the [ color ] sheet from the envelop

_____ Experimenter says, “Please write the definition of each word. You can write either the definition or synonym you practiced. You have 40 seconds for each word. Pencils ready, start.”

_____ Experimenter allowed the total of 3 minutes and 20 seconds.

_____ Experimenter prompted the students to put the [ color ] sheet back to the envelop

_____ Experimenter prompted the students to take out the [ color ] sheets from the envelop

_____ Experimenter said, “Now, please circle the appropriate part of speech and a synonym for each word. You will have 15 seconds for each word. After you are finished with the third sheet, turn the page over and work on another sheet. Please write a new
sentence using each word. Please do not write the definition. I look forward to unique, interesting sentences. You have 1 min. for each word. Of course, you can move on your own pace. Please start.”

_____ Experimenter waited for 6 min. and 15 seconds.

_____ Experimenter prompted the students to put the [color] sheet back in the envelop

_____ Experimenter collected the envelops from the students

_____ Experimenter thanked the students for participation