THE EFFECTIVENESS OF AN INSTRUCTIONAL ASSISTANT LED SUPPLEMENTAL EARLY READING INTERVENTION WITH URBAN KINDERGARTEN STUDENTS

DISSERTATION

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

By
Amanda L. Yurick, M.A.

* * * * *

The Ohio State University
2006

Dissertation Committee:

Dr. Gwendolyn Cartledge, Advisor
Dr. William L. Heward
Dr. Joseph Gliem

Approved by

________________________
Advisor
College of Education
ABSTRACT

The study investigated the effectiveness of a supplemental early reading intervention program on the phonemic awareness and alphabetic principle skills of students identified as being at-risk for reading failure. Six instructional assistants and two graduate students in special education provided small-group instruction for 61 at-risk Kindergarten students across three elementary schools. Thirty-two Kindergarten students from the three schools served as controls. Participants were selected based on low scores on the beginning- and middle-of-year benchmark assessments as measured by the Dynamic Indicators of Basic Literacy Skills (DIBELS). Scores on the letter-word identification and word attack subtests of the Woodcock-Johnson Tests of Achievement-III (WJ-III) were used to verify risk status. Total instructional time varied across instructional assistants. Student participants received between 6.85 and 13.70 hours of instruction across 16 weeks of intervention. Multiple linear regression was used to analyze the effects of participation in instruction, total instructional time, and level of treatment quality on the letter-word identification and word attack subtests of the WJ-III.
Dedicated to my parents
ACKNOWLEDGMENTS

I owe the greatest debt of gratitude to my advisor, mentor, and personal role model, Dr. Gwendolyn Cartledge. She initiated this endeavor by recognizing my potential, she expertly guided and supported me throughout the journey, she taught me how to lead, and she never faltered in modeling the highest level of compassion, integrity, strength, and wisdom. Over these years, she has become more than my advisor. She is my inspiration for passionately pursuing the highest quality of life and education for all students, particularly urban children. If I achieve a fraction of her dream for these children, I will have succeeded.

I also need to thank my committee members. With Dr. Gliem’s expertise in research method and analysis I was able to expand my own research knowledge and experience by attempting and completing this study. Likewise, Dr. Heward has been invaluable to me by teaching me how to conceptualize, analyze, and evaluate educational research. I would like to thank him for always asking the right questions and truly challenging me.

I express my sincere gratitude to my colleagues Shobana Musti-Rao, Ya-yu Lo, Lefki Kourea, Mandy Cooper, Lisa Klett, and Corinne Murphy. Each in your own way has contributed to this study. Mandy and Lisa provided high quality intervention, Lefki was always there to help, and Shobana, and Ya-yu have given me the advice and
guidance that helped me become a better student, teacher, and researcher. I would also like to thank my writing partner, Corinne, for going through this program with me. Finally, the support and true friendship Lis provided helped see me through this project.

I also acknowledge my family’s unending love and humor. My father’s high expectations and affirmation inspires me to become the scholar he believes I can be. My mother’s dry wit and strong will reminds me “how to have a good laugh and a good fight”. My brother has led the way for me and modeled an honest work ethic and determination throughout our lives together. I could not have reached this point in my life without Chris, who is my deepest confidant and whose friendship I hold most dear. And of course I am thankful to Mo. When I needed it most, he was stable and unwavering (to a fault). Like it or not, you will always be my rock (and my PLPRB).
VITA

June 21, 1977 .......................... Born – Akron, OH

1999 ................................. B. M. in music education, Kent State University

1999-2001 ............................. Teacher, ACT 1, inc., Akron, OH

2003 ................................. M.A. in special education, The Ohio State University

2003 – Present ........................ Graduate Research Associate, The Ohio State University

PUBLICATIONS


FIELDS OF STUDY

Major Field: Education

Specialization: Special Education and Applied Behavior Analysis

Minor: Quantitative Research Methods in Human Resource Development
TABLE OF CONTENTS

Abstract ........................................................................................................................................ ii
Dedication ................................................................................................................................. iii
Acknowledgements ................................................................................................................ iv
Vita ............................................................................................................................................ vi
List of Tables ........................................................................................................................... xii
List of Figures .......................................................................................................................... xiv

Chapters:

1. Introduction ......................................................................................................................... 1
   The report of the National Reading Panel ................................................................. 4
   Alphabets ......................................................................................................................... 5
      Phonemic awareness ................................................................................................. 5
      Phonics instruction ..................................................................................................... 8
   Final recommendations of the NRP ........................................................................ 11
   Purpose of the study ....................................................................................................... 14
   Research Questions ......................................................................................................... 14

2. Literature Review ............................................................................................................. 16
   Effective early intervention in reading ................................................................. 16
      Core reading program and the role of supplemental instruction .................. 17
         Basal reading programs ...................................................................................... 20
      Characteristics of effective intervention programs ...................................... 23
         Explicit instruction ................................................................................................. 26
         Intensive instruction ............................................................................................... 26
         Supportive instruction .............................................................................................. 29
   Effective early intervention projects ........................................................................ 29
      Project PRIDE ............................................................................................................ 30
4. Results..............................................................................................................107

Woodcock Johnson tests of achievement ......................................................107
Mean scores .................................................................................................108
Word attack..................................................................................................108
Letter word identification ..........................................................................108
Word attack................................................................................................109
Model 1a: Effects of ERI on word attack gain (non-ELL) .................109
Model 1a: Effects of ERI on word attack gain (ELL) ......................110
Model 2a: Effects of treatment duration and quality on word
attack gain (non-ELL). ...........................................................................111
Model 2a: Effects of treatment duration and quality on word
attack gain (ELL). ............................................................................112
Letter word identification ........................................................................112
Model 1b: Effects of ERI on letter word identification gain
( non-ELL) ....................................................................................112
Model 1b: Effects of ERI on letter word identification gain
(ELL) ............................................................................................113
Model 2b: Effects of treatment duration and quality on letter
word identification gain (non-ELL)..................................................114
Model 2b: Effects of treatment duration and quality on letter
word identification gain (ELL). ...................................................115
DIBELS benchmark assessments ..............................................................116
Non-ELL treatment students ..................................................................116
ELL treatment students .........................................................................117
Control students .....................................................................................118
General results ........................................................................................119
Treatment quality ........................................................................................................ 119
  D.C. (School 1) ........................................................................................................ 119
  D.M. (School 1) ........................................................................................................ 120
  A.H. (School 2) ........................................................................................................ 122
  L.P. (School 2) ........................................................................................................ 123
  L.K. (School 2) ........................................................................................................ 124
  R.O. (School 3) ........................................................................................................ 126
  S.W. (School 3) ........................................................................................................ 127
  M.C. (School 3) ........................................................................................................ 128
Social validity ............................................................................................................. 130
  Principal 1 ............................................................................................................. 131
  Principal 2 ............................................................................................................. 131
  Principal 3 ............................................................................................................. 132
  School personnel .................................................................................................... 132
  D.C. (School 1) ........................................................................................................ 133
  D.M. (School 1) ........................................................................................................ 134
  A.H. (School 2) ........................................................................................................ 134
  L.P. (School 2) ........................................................................................................ 134
  L.K. (School 2) ........................................................................................................ 134
  R.O. (School 3) ........................................................................................................ 134
  S.W. (School 3) ........................................................................................................ 134
  M.C. (School 3) ........................................................................................................ 135
Non-ELL treatment students ....................................................................................... 136
ELL treatment students .............................................................................................. 139

5. Discussion .................................................................................................................. 141
  Research question one .............................................................................................. 142
  Research question two .............................................................................................. 145
  Research question three ............................................................................................ 149
  Research question four .............................................................................................. 151
    Instructors ............................................................................................................... 152
    School principals and related personnel .................................................................... 153
  Research question five .............................................................................................. 154
Limitations of the study ............................................................................................... 155
  Training and support ................................................................................................. 155
  Consistency ................................................................................................................ 156
  Specification error ..................................................................................................... 158
  Sample ....................................................................................................................... 160
Implications for practice .............................................................................................. 160
  English language learners ......................................................................................... 160
  Instructional assistants ............................................................................................... 162
Directions for future research ..................................................................................... 164
Summary ....................................................................................................................... 166

xi
References............................................................................................................168

Appendices...........................................................................................................179

A. Parent letter for participation (English) ..............................................179
B. Parent consent for participation in educational research (English)....182
C. Parent letter for participation (Somalian)............................................184
D. Parent consent for participation in educational research
(Somalian)............................................................................................187
E. Parent letter for participation (Spanish) ..............................................189
F. Parent consent for participation in educational research (Spanish)....192
G. Principal letter of agreement ..............................................................195
H. Sample ERI lesson ..............................................................................197
   Lesson Overview ..................................................................................198
   Activity 1 ...........................................................................................199
   Activity 2 ...........................................................................................200
   Activity 3 ...........................................................................................202
   Activity 4 ...........................................................................................203
   Activity 5 ...........................................................................................204
   Activity 6 ...........................................................................................205
   Activity 7 ...........................................................................................206
I. Treatment quality checklist ...............................................................208
J. Daily time log ......................................................................................210
K. Social validity questionnaire (instructors and principals)............212
L. Social validity questionnaire (students) .............................................215
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Demographic information and pretest scores for treatment and control students</td>
<td>77</td>
</tr>
<tr>
<td>3.2</td>
<td>Background information for the instructors</td>
<td>80</td>
</tr>
<tr>
<td>3.3</td>
<td>Mean treatment quality scores by group</td>
<td>93</td>
</tr>
<tr>
<td>3.4</td>
<td>ERI training schedule outline</td>
<td>96</td>
</tr>
<tr>
<td>4.1</td>
<td>Descriptive data for the effects of participation in ERI on word attack gain scores for non-ELL and ELL treatment students</td>
<td>110</td>
</tr>
<tr>
<td>4.2</td>
<td>Coefficients for the effects of participation in ERI on word attack gain scores for non-ELL and ELL treatment students</td>
<td>110</td>
</tr>
<tr>
<td>4.3</td>
<td>Descriptive data for the effects of treatment duration and treatment quality on word attack gain scores for non-ELL and ELL treatment students</td>
<td>111</td>
</tr>
<tr>
<td>4.4</td>
<td>Coefficients for the effects of treatment duration and treatment quality on word attack gain scores for non-ELL and ELL treatment students</td>
<td>112</td>
</tr>
<tr>
<td>4.5</td>
<td>Descriptive data for the effects of participation in ERI on letter word identification gain scores for non-ELL and ELL treatment students</td>
<td>113</td>
</tr>
<tr>
<td>4.6</td>
<td>Coefficients for the effects of participation in ERI on letter word identification gain scores for non-ELL and ELL treatment students</td>
<td>113</td>
</tr>
<tr>
<td>4.7</td>
<td>Descriptive data for the effects of treatment duration and treatment quality on letter word identification gain scores for non-ELL and ELL treatment students</td>
<td>115</td>
</tr>
<tr>
<td>4.8</td>
<td>Coefficients for the effects of treatment duration and treatment quality on letter word identification gain scores for non-ELL and ELL treatment students</td>
<td>115</td>
</tr>
</tbody>
</table>
4.9: Results of the social validity questionnaire for school principals, data manager (School 2) and Reading First coordinator (School 3) ........................................133

4.10: Results of the social validity questionnaire for the instructors ........................................136

4.11: Summary of non-ELL treatment student’s social validity questionnaire responses ..........................................................................................................................138

4.12: Summary of ELL treatment students’ social validity questionnaire responses ...140
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Slope of mean pretest and posttest word attack scores for non-ELL treatment, ELL treatment, and control students</td>
<td>108</td>
</tr>
<tr>
<td>4.2</td>
<td>Slope of mean pretest and posttest letter word identification scores for non-ELL treatment, ELL treatment, and control students</td>
<td>109</td>
</tr>
<tr>
<td>4.3</td>
<td>Proportion of non-ELL treatment students with intensive, strategic, and benchmark intervention recommendations in fall and spring</td>
<td>117</td>
</tr>
<tr>
<td>4.4</td>
<td>Proportion of ELL treatment students with intensive, strategic, and benchmark intervention recommendations in fall and spring</td>
<td>118</td>
</tr>
<tr>
<td>4.5</td>
<td>Proportion of control students with intensive, strategic, and benchmark intervention recommendations in fall and spring</td>
<td>119</td>
</tr>
<tr>
<td>4.6</td>
<td>Treatment quality scores for D.C. with a mean line</td>
<td>120</td>
</tr>
<tr>
<td>4.7</td>
<td>Treatment quality scores for D.M. with a mean line</td>
<td>121</td>
</tr>
<tr>
<td>4.8</td>
<td>School 1 instructors’ mean percentage treatment quality, duration, and student gain on LWID and WA posttest</td>
<td>122</td>
</tr>
<tr>
<td>4.9</td>
<td>Treatment quality scores for A.H. with a mean line</td>
<td>123</td>
</tr>
<tr>
<td>4.10</td>
<td>Treatment quality scores for L.P. with a mean line</td>
<td>124</td>
</tr>
<tr>
<td>4.11</td>
<td>Treatment quality scores for L.K. with a mean line</td>
<td>125</td>
</tr>
<tr>
<td>4.12</td>
<td>School 2 instructors’ mean percentage treatment quality, duration, and student gain on LWID and WA posttest</td>
<td>126</td>
</tr>
<tr>
<td>4.13</td>
<td>Treatment quality scores for R.O. with a mean line</td>
<td>127</td>
</tr>
<tr>
<td>4.14</td>
<td>Treatment quality scores for S.W. with a mean line</td>
<td>128</td>
</tr>
</tbody>
</table>
4.15: Treatment quality scores for M.C. with a mean line ................................................. 129

4.16: School 3 instructors’ mean percentage treatment quality, duration, and student gain on LWID and WA posttest ................................................................. 130
CHAPTER 1

INTRODUCTION

Underachievement in reading is a well-documented and persistent problem. The National Center for Educational Statistics (NCES, 2003) has reported that 38% of all fourth-grade students scored below the “basic level” in reading. Students in urban areas fared worse, with 52% reading below the basic level. This means that more than half of urban fourth-grade students do not have the reading skills necessary to perform grade-level work. Students attending school in high poverty urban areas are at the greatest risk (Foorman & Moats, 2004). This population may have failure rates in reading as high as 70% (Good, Gruba, & Kaminski, 2001). Furthermore, the National Reading Panel (NRP, 2000) has indicated that at least 20% of children demonstrate reading deficits before the third grade. The third grade is a critical benchmark because a longitudinal study conducted by Felton and Pepper (1995) showed that many children, who do not acquire sufficient word recognition skills by the third grade, rarely make up these deficits. It is obvious that, even though reading has gained national attention, a substantial number of students are still not benefiting from effective instruction and intervention (Bursuck et al., 2004).
Reading achievement in urban schools deserves special attention since these children exhibit the most risk. Students in urban districts are more likely than their suburban peers to be poor, represent a minority group, and have parents with little or no postsecondary education. Recent data from the NCES (2003) indicate that, “At the start of Kindergarten in both reading and mathematics, black children had lower mean achievement scores than other racial/ethnic groups, and children with family risk factors had lower achievement scores than their peers with fewer risk factors. These achievement gaps grew wider from the start of Kindergarten in fall 1998 to the end of 3rd grade in spring 2002” As stated previously, reading achievement in the third grade appears to be a tipping point for the prognosis for future educational accomplishments, so these data underscore the point that urban minority children with family risk factors must receive early intensive intervention efforts. The NCES cites work by West, Denton, and Germino Hausken (2000) that indicates that white children tend to score higher on letter recognition and beginning and ending sound isolation tasks than do their Hispanic and black peers. These conclusions suggest that early reading intervention should aim to close the achievement gap between white and minority children on alphabetic and phonemic awareness tasks.

The long-term consequences of reading failure include risk for behavioral problems, greater probability of being referred for special education, delinquency, unemployment, dropping out of school, and even incarceration. The most current data available for drop out rates in the United States is for the year 2000. In that year approximately 5% of students age 15-24 failed to complete a high school degree and that figure has remained stable since 1987. However, in large cities such as Los Angeles,
Chicago, Boston, and New York, the drop out epidemic may be even more pronounced, with estimates between 40-60%. In fact, according to the NCES, in the year 2000, “young adults living in families with incomes in the lowest 20 percent of all family incomes were six times as likely as their peers from families in the top 20 percent of the income distribution to drop out of high school”. The impact of under-educating American students cannot be overstated. Clifford (1990) likened dropping out of high school to educational suicide, and made the point that this systematic failure of the educational establishment to meet the needs of our students and ensure their graduation affects students of all abilities and potential. The loss of these students’ potential threatens the economical, intellectual, cultural, and social well being of the future of our country. The link between student achievement and the productivity and health of our future labor market and economy is obvious (Siegle & McCoach, 2002; Office of Educational Research and Improvement, 2000). Since literacy is the key to all subsequent learning and achievement, it is critical that young children receive the most effective reading instruction possible.

To further compound the task of providing effective early reading instruction for all students, the increasing levels if diversity found in many urban schools contributes the unique challenge of educating children from all backgrounds, including students whose primary language is not English. A multitude of factors such as poverty, low education levels of parents, poor teacher quality, and inadequate instructional methods and materials threaten to diminish educational outcomes for these students (Snow, Burns, & Griffin, 1998). English language learners (ELLs) may be particularly at risk since they may come to school with little or no receptive or expressive English language skills.
Since these students often experience reading difficulties in early grades, many of them continue to trail behind classmates throughout the elementary years and may never catch up with their peers (Ortiz & Graves, 2001). The resulting effect is sometimes referred to as the Matthew Effect, which contends that the rich continue to get richer, while the poor get poorer (Stanovich, 1986). We must determine if the principles that guide the development of early reading instruction for English speaking students can be effective for ELLs too. Otherwise, these students risk continued disproportionate referral for special education services, thus cementing the Matthew Effect through a diminished educational trajectory and poor prognosis for future schooling.

The National Reading Panel Report

In 1997, a panel of 14 members, along with subgroup committee members, was charged with the task of identifying, clarifying, and synthesizing the most effective components of reading instruction. Specifically, the members of the NRP were charged with analyzing, compiling, and disseminating usable findings from the research in a timely fashion. First, the NRP delineated the basic areas of reading instruction to be studied: (a) Alphabetics, which included phonemic awareness and phonics instruction; (b) Fluency, the ability to read with speed and accuracy; and (c) Comprehension, which included vocabulary instruction and text comprehension. Additionally, the NRP chose to examine teacher education and reading instruction, and computer technology and reading instruction. However, for the purpose of the present study, only those areas that pertain directly to early reading instruction will be examined, specifically alphabetics, which includes phonemic awareness instruction and phonics instruction.
Alphabetic principle is demonstrated when the learner has the ability to associate sounds (phonemes) with letters (graphemes), and can apply this association to read words (Moats, 1999). There are two instructional areas that address the skills necessary to demonstrate knowledge of the alphabetic principle and the NRP report appropriately focused on them: phonemic awareness and phonics. The point should be made that phonemic awareness and phonics are not synonymous terms. Phonemic awareness deals with the isolation of sounds in spoken words. Phonics entails the depiction of grapheme-phoneme (i.e. letter-sound) correspondences to decode and spell words. Although a clear relation exists between phonics and phonemic awareness, they should be discussed separately.

Phonemic awareness. Although it may seem counterintuitive, there are actually more phonemes in the English language than there are letters in the alphabet. There are between 41 and 44 identified phonemes (Snow, Burns, & Griffin, 1998) which consist of individual letter sounds and special sounds created by letter combinations, such as /sh/. Phonemic awareness is the ability to isolate and manipulate phonemes, which are the smallest units in spoken language, in spoken words. For example, a student who responds to the prompt, “What is the first sound in paste?” with the sound /p/ has demonstrated some level of phonemic awareness. The above example illustrates only one phonemic awareness task, phoneme isolation. Typically, students will be assessed on phoneme identity, categorization, blending, segmentation, and deletion. Each of these tasks requires the student to manipulate the sounds of the words in different ways. In identity tasks, students are asked to provide the common sound of some given words (as in boy,
bike, and bell). In phoneme categorization, students must choose the word that does not fit the sound pattern (such as in bus, bun, rug; rug would be the inappropriate word). In a phoneme-blending task, the student is provided a word by the instructor in segments. The students must then blend the segments together to provide a smooth verbalization of the word. If the instructor says, “What word is /b/ /u/ /g/ /s/?” the students should respond, “bugs”, for example. Phoneme segmentation tasks are the opposite of blending tasks. The instructor says the blended word to the student who is supposed to respond with each phoneme in segmentation. For example, the instructor says, “How many phonemes are in ship?” The student would respond that there are three: /sh/ /i/ /p/. In a phoneme deletion task, the student is asked to delete one phoneme from a blended word. For example, the instructors would ask, “what is smile without the /s/?” The correct response would be mile.

In the NRP’s meta-analysis of the research literature on phonemic awareness, 52 of a potential 1,962 articles qualified under their criteria, which included the following:

1. The design of the study had to be either true experimental with a control group or quasi-experimental in a multiple baseline design.
2. The study was published in a peer-reviewed journal.
3. The directional hypothesis was that phonemic awareness improves reading performance over other forms of instruction or no instruction.
4. The phonemic awareness training independent variable in the study was not confounded by other instructional methods and activities.
5. The study needed to report statistics that would permit the calculation of an effect size.
An estimate of “effect size” is particularly relevant in educational research since it provides an estimation of practical importance, not only a negation that the result occurred by chance, as other statistical tools do. An effect size of 1.0 indicates that the treatment group performed one full standard deviation above the control group on average. An effect size of 0 indicates that the treatment group and the control group means are identical. The standard judgment is that effect sizes of .80 and greater are considered large, .50-.79 moderate, and .20-.49 small. Effect sizes of 0-.19 are considered too small to be relevant. The NRP calculated three effect sizes for each study in the meta-analysis: phonemic awareness; reading; and spelling.

The purpose of the meta-analysis was to examine how phonemic awareness instruction affected children’s phonemic awareness, reading, and spelling performance. The results indicated that the overall effect size for phonemic awareness outcomes was large, .86. The overall effect sizes for reading and spelling were .53 and .59, respectively, which qualifies as moderate. The NRP concluded that phonemic awareness instruction is an effective tool to teach children to attend to and manipulate phonemes. Furthermore, the NRP definitively stated that the ability to manipulate phonemes in speech improves children’s reading. It should also be noted that although some argue that phonemic awareness instruction ignores reading comprehension, phonemic awareness training is shown to have an effect on children’s overall comprehension.
Although less robust than effect sizes on word reading in isolation, the NRP reasons that this boost in reading comprehension is intuitive since reading comprehension is predicated on effective word reading. Finally, the NRP concluded that phonemic awareness instruction was effective in helping children learn how to spell. More importantly, these effects of the instruction on spelling were durable over time.

The NRP examined the results of the meta-analysis for moderators on the effectiveness of phonemic awareness instruction. That is, did certain characteristics of the instruction or conditions produce larger effect sizes? A particularly strong moderator on performance was the explicitness and focus of the instruction. Generally, as the instruction became more explicit, the effect sizes were larger. The duration of intervention also exerted influence on the outcomes. Interestingly, the longest treatment protocols were not the most effective. The treatment duration that produced the largest effect sizes was between 5 and 18 hours of instruction. Treatment durations that were shorter than 5 hours or longer than 18 produced smaller effect sizes. Finally, it is notable that studies including students in preschool and Kindergarten demonstrated larger effect sizes than those including students in first grade and above. The synthesis of these findings indicates that a phonemic awareness intervention for young children that is explicit in nature and lasts between 5 and 18 hours total shows the most promise for improving reading outcomes.

Phonics instruction. Several types of phonics instruction exist. Some of these approaches include: (a) Synthetic phonics, in which children are taught to convert letters to sounds or phonemes followed by blending into whole words; (b) Analytic phonics, which de-emphasizes the isolation of sounds and focuses on analyzing letter-sound
relations in whole words; (c) Phonics in context is a type of phonics program that uses letter-sound relations and context cues to identify unknown words in texts; (d) Analogy phonics uses known parts of words to identify new words. Although differences among these types of phonics programs exist, they all share a systematic element, in which a planned set of phonic elements are taught sequentially and explicitly. Furthermore, the overarching goals of all phonics programs is to help children acquire workable knowledge of the alphabetic code in order to decode unknown words, recognize familiar words accurately and automatically, and comprehend text.

The inclusion criteria for the NRP’s meta-analysis of the phonics literature was identical to the criteria for the phonemic awareness meta-analysis with the additional stipulation that studies used for the phonemic awareness meta-analysis could not be included in the present analysis. Seventy-five studies were screened and 38 were retained, from which 66 treatment-control comparisons were drawn. Effect sizes were calculated on six outcome measures: (a) Decoding of regularly spelled words; (b) Reading novel pseudowords; (c) Reading irregularly spelled miscellaneous words; (d) Spelling words; (e) Reading comprehension; and (f) Accuracy of oral reading. Although the studies were examined for these 6 outcomes, few reported all six. Most (75%) of the effect sizes calculated were on reading and spelling single words, while 24% accounted for text reading.

The overall effect size for phonics instruction was .44. Three categories of phonics instruction (synthetics, larger-unit, and miscellaneous types) were analyzed separately and yielded effect sizes of .45, .34, and .27, respectively. The NRP also analyzed the data according to the delivery system used: individual tutoring, small
groups; and whole class. All three methods were effective, with effect sizes of .57, .43, and .39, respectively. The results indicated that phonics instruction was optimal at the Kindergarten and first grade levels. Effect sizes for these two groups were .56 and .54, and the effect size for studies using 2nd through 6th graders was .27. Phonics instruction was shown to be beneficial for all students, but particularly strong effect sizes were reported for at-risk Kindergarteners and at-risk first-graders. Studies including students from these two groups demonstrated effect sizes of .58 and .74. Studies including students with specific reading disabilities yielded an effect size of .32. Finally, the NRP noted that phonics instruction produced significantly greater growth in reading comprehension and spelling than non-phonics instruction. The effect size for comprehension was .51 and .67 for spelling. These results were relevant for students in Kindergarten and first grade, but not for students in 2nd grade and above.

The NRP concluded that systematic phonics is an effective tool to teach children to read and should be used as a part of a successful reading program. However, the NRP cautions that phonics in isolation without the benefit of application to context is unlikely to be effective and those programs that focus solely on sound isolation should not be adopted. The NRP went on to advise that teachers’ roles need to be considered. That is, although scripted phonics programs produced the greatest effect sizes, the elimination of teacher judgment may reduce the teacher’s interest and motivation to teach such a program. However, it should be noted that there is no empirical evidence that well-constructed scripted programs eliminate teacher judgment. The misconception that scripted programs strip teachers of a critical role in instructional delivery may be a major obstacle to implementing these effective programs. The conclusion drawn from this point
is that future research and development of phonics programs should examine how to maintain instructional consistence while encouraging the teacher’s unique contributions. Clearly phonics instruction is an empirically validated effective practice, however, the consumer’s (in this case, the teacher’s) dissatisfaction with the delivery method speaks to the treatment’s level of social acceptability and therefore merits future research into consumer satisfaction, not necessarily treatment utility and effectiveness.

Final recommendations of the NRP

The NRP definitively endorsed the implementation of both phonemic awareness and phonics instruction as part of a core reading program. However, the NRP emphatically stated that neither phonemic awareness nor phonics alone or in tandem constitute a comprehensive reading program. Acquiring phonemic awareness and workable knowledge of the alphabetic code is a means rather than an end. These areas of instruction are most appropriately emphasized at the early stages of reading and their proportional representation in the comprehensive reading program should fade gradually, particularly after second grade.
Future research in this area should address teacher training in and motivation to teach phonemic awareness and phonics. The NRP cited research that raised questions regarding the teacher’s level of phonemic awareness because some studies have shown that the teachers do not accurately manipulate the phonemes. Obviously, this skill deficiency would negatively impact the instructional effectiveness of any phonemic awareness program, particularly unscripted ones. Unfortunately, only one third of the studies in the NRP’s meta-analysis reported data on treatment fidelity, the extent to which the implementer of the intervention adheres to stated guidelines of the protocol. Compromised treatment fidelity (also known as procedural fidelity and treatment integrity) undermines the credibility of the study’s findings and should be not only included in research reports, but also may merit more detailed study of the effects of fidelity on treatment outcomes.

When considering the lack of teachers’ phonemic awareness skills, an unfortunate conundrum arises. Scripted phonics and phonemic awareness programs consistently demonstrated stronger effect sizes than unscripted or more ambiguous programs. Scripts ensure the standardization, consistency, and fidelity of the instructional program, and can mitigate any teacher deficiency in this skill area. However, as the NRP pointed out, many teachers are uncomfortable with scripts and prefer programs that enable more unique teacher contributions and participation. Whether using a scripted or an unscripted curriculum, teachers need to be highly skilled and well trained in the delivery of an effective core reading program (Snow, et al., 1998).
Ideally, a well-trained teacher using evidenced-based curricula would be preferred to provide supplemental reading interventions for students who struggle in the core program and require more explicit and intensive instruction to be successful. However, many teachers do not have the time or resources to conduct such interventions. This is particularly true for urban settings in which many students need such assistance. In these circumstances, paraprofessionals, also known as aides or instructional assistants, can help provide supplementary intervention. Several studies have found that when instructional assistants are well trained and receive feedback and supervision, students’ reading can benefit from the instruction they provide (Bursuck et al., 2004; Mercer, Campbell, Miller, Mercer, & Lane, 2001). However, future research is needed to investigate methods for using instructional assistants to deliver intervention. Additionally, we need research on preserving program fidelity and effectiveness, whether through scripting or other means, while increasing teacher satisfaction so that these evidence-based practices can be unequivocally adopted in our schools.
Purpose of the Study

The purpose of the study was to extend previous research by Musti-Rao (2005), examining the effects of the Early Reading Intervention (ERI; Simmons & Kame'enui, 2004) program on at-risk Kindergarten students’ phonemic awareness skills. Specifically, this study implemented the ERI with 61 experimental students across three high-poverty, urban elementary schools and employed six Instructional Assistants (IA’S) and two graduate students to deliver the intervention. Twenty-seven students in the experimental group were Somalian or Latino students, also known as English Language Learners (ELLs). The primary goal of this study was to determine the amount of variance in outcome scores for both ELL and non-ELL students that could be attributed to participation in the ERI, the duration of the treatment, and the level of integrity and quality with which the intervention was delivered.

Research Questions

The current study addressed the following research questions:

1. How much variance in student gain scores as measured by the letter-word identification (LWID) and word attack (WA) subtests from the Woodcock-Johnson Tests of Achievement III (WJ III) can be explained by participation in the ERI for both ELL and non-ELL students?

2. How much variance in student gain scores as measured by LWID and WA subtests from the WJ III can be explained by treatment duration and treatment quality for both ELL and non-ELL students?

3. What proportion of ELL and non-ELL treatment students will improve their DIBELS recommendations?
4. As a measure of social validity, to what extent will instructors, school principals, and related personnel view the supplemental instruction as beneficial to their students?

5. To what extent will students rate the supplemental instruction as beneficial?
This chapter reviews the literature on (a) characteristics of effective early intervention in reading, particularly as it relates to students at risk for reading and related disabilities, (b) the acquisition of phonological awareness and other early reading skills for English Language Learners (ELLs) and; (c) critical aspects of intervention delivery. Specifically, the roles of treatment integrity and duration of treatment in affecting students’ outcome scores will be explored.

Effective Early Reading Instruction

For many children reading difficulties can be prevented with good instructional strategies and materials in a whole-class format (Snow, Burns, Griffin, 1998). In fact, according to estimates from the three-tiered intervention model (Kamps & Greenwood, 2005), approximately 80% of students will be successful with effective regular classroom instruction. The second tier of approximately 15% of students will require a supplemental intervention that enhances the primary curriculum in order to read successfully. Finally, the third tier consists of interventions that are “specifically designed and customized for students with marked difficulties in reading or reading disabilities, and who have not
responded to primary and secondary prevention efforts” (Kame’enui, Good, Simmons, & Chard, 2002, p. 4). Students who do not respond to regular classroom instruction or some type of supplemental intervention efforts are likely to be referred for special education due to a specific reading disability. This is especially troublesome since phonological processing deficits are the most common cause of reading disabilities (Torgesen, Wagner, & Rashotte, 1997a) and 85% of all special education referrals are determined to be eligible for services (Gottlieb, Alter, Gottlieb, & Wishner, 1994). Some estimates suggest that 10% to 20% of school age children have been diagnosed with specific reading disabilities (Harris, & Sipay, 1990), while others have placed the figure even higher at 30%, though this figure may be somewhat inflated (Vellutino & Scanlon, 2002).

Core Reading Programs and the Role of Supplemental Instruction

As part of a longitudinal study of 1407 entering Kindergarten students, Vellutino et al. (1996) hypothesized that early reading impairments are caused by limitations in early reading experiences and/or classroom instruction. Specifically, the authors anticipated that the type of instruction found in the Kindergarten classroom would be significantly correlated to reading achievement in the first grade. Additionally, it was proposed that a substantial proportion of poor readers could be remediated to within normal ranges using an early, intensive one-on-one tutoring intervention. Poor readers were initially identified by the classroom teacher and that status was confirmed by a score at or below the 15th percentile on the Word Identification or Word Attack subtests of the Woodcock Johnson Reading Mastery Tests – Revised (WRMT-R; Woodcock, 1987). Additionally, in order to qualify, students needed to have an I.Q. of at least 90 on the Wechsler Intelligence Scale for Children – Revised (Wechsler, 1974). Finally, students
with sensory, physical, and/or neurological impairments, emotional disorders, or frequent school absences were excluded from the study. These criteria resulted in a selection of 118 impaired readers from the larger sample, which accounted for approximately 9% of the population originally evaluated. Seventy-six of these impaired readers were randomly assigned to receive a one-on-one tutoring intervention for 30 minutes daily through their first-grade year and during the fall of their second-grade year if they were found to still have deficits at the end of first-grade. The remaining 42 students in the impaired group served as controls and received whatever form of remediation their school offered. Most students received small-group instruction (n=26), but some (n=16) received one-on-one tutoring two or three times per week. Additionally, a comparison group of 65 normal readers (45th percentile of the WRMT-R; I.Q. ≥ 90) was used. The treatment was an individual tutoring program geared toward students’ strengths and weaknesses and included a balance of word-level (e.g., phonological decoding and word identification) and text processing skills (e.g. reading connected text and comprehension activities). Fifteen minutes of the intervention period was devoted to fluency and connected reading and 15 minutes was devoted to phonemic awareness, sight words, and writing.

Results of the intervention portion of the study indicated that most of the impaired reader group (67%) was reading within the normal range after only one semester of intervention. Three percent were still below average and only 1.5% continued to be severely impaired, which is a substantial decrease from 9% at the outset of the study. Almost 45% of impaired students in the control group who received tutoring two to three times per week scored within the normal range and 19% of the students receiving small group instruction fell within this range. However, 25% of the control students receiving
two to three sessions of individual tutoring per week scored below the 15th percentile. This figure indicates that although many students can benefit from lesser amounts of tutoring, more frequent sessions may be necessary to remediate students who score at the lowest levels. Thus, although 9% of the students identified at the outset of the study could have been identified with a reading disability, only 1.5% of the sample of the impaired readers who received the daily intervention would have qualified for this diagnosis.

Another objective of this study was to determine the relationship between the type of instruction offered in the Kindergarten classroom and later reading achievement. Specifically, the authors noted that both code-based (e.g. phonemic awareness and phonics) activities and text-based (e.g. shared reading) activities differentially affected first-grade reading achievement. Upon entry into Kindergarten students were divided into either the “high letter identification” (HLID) group or a “low letter identification” (LLID) group based on scoring either above or below the 40th percentile on the Letter Identification subtest of the WRMT-R. Then in first grade, these students were rated on a five-point scale of general reading achievement by their teachers. In the Kindergarten year, teachers of LLID students tended to spend more time on phonemic awareness activities than teachers of HLID students. Consequently, children from the LLID group who were rated average or better than average readers by their first-grade teachers were more likely to come from Kindergarten classrooms that focused on phonemic awareness activities. First-grade HLID children who were rated average or better than average readers were more likely to come from the Kindergarten classrooms that stressed shared reading activities such as teacher-led reading from large-print books. This finding suggests that children from the LLID group may have entered Kindergarten without the
requisite phonemic awareness skills necessary to benefit from text-based shared reading activities, while the students from the HLID already mastered the requisite phonemic awareness skills necessary to benefit from this type of instruction. It is therefore necessary that children’s entry skills be used as determinants of the type and intensity of both regular classroom instruction and supplemental interventions. Children who enter Kindergarten without substantial phonemic awareness skills need reading instruction that focuses on these skills before text-based instruction begins.

In general, the authors of the study concluded that the current incidence of reading disabilities is grossly overestimated. Most students can become functional readers with early and intensive intervention that is responsive to individual strengths and weaknesses. Further, and perhaps most importantly, the findings from this study support the contention that, “reading difficulties in beginning readers are caused by experiential and instructional deficits rather than by neurological deficits” (Vellutino & Scanlon, 2002, p. 317). This conclusion calls attention to the importance of identifying early reading difficulties and addressing them directly within the regular classroom instruction. Labeling and removing children from the classroom for specialized instruction suggests that the problem is within the child and obscures from the teacher the possibility that something is wrong with the classroom instruction (Strickland, 2002). The earliest attempts at intervention for reading difficulties should take place within the regular classroom, within the regular core-reading program.

*Basal reading programs.* The core reading, or basal, program is at the center of every student’s literacy program (Strickland, 2002). Yet, very few studies have empirically examined the effectiveness and quality of these programs and fewer still have
established the efficacy of these programs for at-risk students. Even more disturbing is that fact that even though converging evidence supports the predictive validity of phonological awareness on later reading achievement (Adams, 1990; Barnett, 2002; Vellutino & Scanlon, 2002), there is only one published study that evaluates the level at which this critical skill is addressed in basal reading programs. Since deficits in phonological awareness, or the ability to manipulate sounds in words, has been identified as a cause of reading disabilities (Smith et al., 2001; Lyon & Chhabra, 1996) it is critical that basal reading programs adequately teach this skill as a means of preventing future reading disabilities.

Smith and colleagues examined four basal reading programs: (a) Harcourt Brace Jovanovich’s The Treasure Tree (Farr et al., 1993), (b) Houghton Mifflin’s Level K1 All About Me and Level K2’s Let’s Be Friends (Pikulski et al., 1991, 1992), (c) Macmillan/McGraw Hill’s Tell a Story/Sing a Song (Aoki et al., 1993), and Scott Foresman’s Celebrate Reading (Allington et al., 1993). The programs were coded for (a) presence or absence of any phonological awareness activities, (b) dimensions of phonological awareness, (c) the conspicuousness (or explicitness) of the strategies, (d) the presence or absence of scaffolds and supports to mediate student learning, and (e) the number of instructional practice examples. The results of this study shed light on substantial deficits in basal reading programs. Although all of the programs included activities that were specifically intended to develop phonological awareness, the quantity and quality of examples that targeted phoneme segmentation and blending, the activities most highly correlated with reading achievement, were insufficient to produce gains in reading. Furthermore, the instruction was not conspicuous enough to call attention to the
features of words that children need to learn. For example, children were often encouraged to listen to a teacher model and note the sameness of certain features. In fact, none of the programs requested the teacher to explicitly model the critical features that the children were being asked to note. Another troubling finding was that all of the programs required student responses only at the word level, rather than the phoneme level, which is in direct conflict with the most current evidence of effective literacy instruction. Reading research clearly indicates that good readers cannot only identify phonemes in words, but can also manipulate phonemes in the initial and final positions. This type of instruction was virtually absent in all the curricula examined.

Based on their research, Smith and colleagues succinctly summarize years of converging research findings on phonological awareness and provide the following guidelines: (a) all children can benefit from early phonological awareness instruction prior to the onset of formal reading instruction, (b) some dimensions of phonological awareness are more crucial to reading achievement than others (e.g. blending and segmenting), (c) conspicuous or explicit teaching strategies promote phonological development in children, and (d) features such as the size of the phonological unit, position of the phoneme, and phonological properties require scaffolding. Considering these guidelines, it is clear that these four basal reading programs have recognized the importance of phonological awareness but have not incorporated the necessary elements of effective instructional design to impart this knowledge to students. The authors conclude by making the following recommendations to publishers (p. 48):
1. Include instruction in the “big ideas” of blending and segmenting as auditory tasks-particularly at the phoneme level.

2. Systematically sequence tasks.

3. Provide examples of explicit instruction to increase the conspicuousness of strategies.

4. Increase opportunities to produce sounds at the phoneme level.

5. Provide generous suggestions for scaffolding of tasks, materials, and the amount of teacher assistance.

When well known instructional principles such as systematically sequencing tasks, providing multiple opportunities to practice phoneme level skills such as blending and segmenting, and sufficiently scaffolding difficult tasks are applied to early reading instruction and intervention, significant improvements in the rate of reading growth for students with many types of reading problems can be achieved (Torgesen, Wagner, & Rashotte, 1997a).

**Characteristics of Effective Intervention Programs**

The guidelines for effective instruction outlined above serve as a foundation for all core reading and supplemental programs. However, for students most at risk for reading failure, these elements need to be made more direct, explicit, intensive, and supportive in small groups (Foorman & Torgesen, 2001). We know that students must demonstrate phonological awareness and mastery of the alphabetic principle in order to be successful readers. However, what researchers and educators don’t agree on is how that mastery is to be attained. How direct, explicit, intensive, and supportive does the instruction need to be? As illustrated by the previous review of some common basal
programs, the gap between the intervention research knowledge base and practice is evident. Not only are basals lacking in many of the principles of effective instruction, they are often built upon a shaky ground of broad assumptions such as, “that children will learn to spell the words they learn to read, that they will learn vocabulary by highlighting the story words prior to reading passages, and that integration of vocabulary, spelling, and writing will occur automatically” (Foorman & Torgesen, 2001, p. 205).

A recent study conducted by Foorman and colleagues (1998) examined the effects of three reading programs of varying levels of directness on the word reading skills of 285 first- and second-grade students across eight Title I schools in the Houston area. The three levels of the independent variable were (a) implicit code, (b) embedded code, and (c) direct code. Implicit code instruction adheres to the whole language framework and gives priority to children’s self-constructed meaning. The teacher is considered the facilitator of instruction, rather than the director, and any phonics-based activities are conducted opportunistically. The embedded code instructional strategy used somewhat structured activities in sound-spelling patterns from trade books. Typically, the phonics activities occur at the level of onset-rime analyses, rather than the phoneme level, which is more sensitive. Direct code instruction is characterized by the use of controlled vocabulary texts, in which students are directly taught letter-sound correspondences. First, the instruction focuses on developing a base of print conventions and knowledge using language games and rebus activities. Next, the students are taught to sound out unknown words during independent reading activities using a graduated series of books. Finally, students work to develop reading strategies, practice spelling, writing, and using
their phonics knowledge by reading trade books. Out of the 66 classrooms in the study, 19 used the implicit code approach, 20 used embedded code, and 14 used the direct code.

Results indicated that after controlling for differences in age, ethnicity, and initial level of phonemic awareness, children taught by the direct code approach improved their word reading at a faster rate and obtained higher posttest scores than students in the implicit code group. The direct code and embedded code group did not differ on average, considerable individual differences emerged among the implicit code and embedded code groups. A sizeable percentage of students from the implicit- and embedded-code groups (46% and 44% respectively) did not show any measurable improvement in word reading, averaging a growth rate of 2.5 per year or less. In comparison, only 16% of students from the direct code group fell into this category. In fact, by the end of the study, students from the direct code group were more or less normally distributed. It is worth noting that the scores of the embedded code group tended to fall in between the scores of the direct code and implicit code students, which suggests that an overall gradient of generalizability emerges with direct code students representing the highest scores, embedded code students demonstrating somewhat lesser gains, and implicit code students falling into the lowest scoring category.

In a 2001 paper, Torgesen and Foorman outlined three broad characteristics necessary for effective intervention. The authors recommend that effective interventions are: (a) more explicit, (b) more intensive, and (c) more supportive than the regular classroom instruction. Each of these characteristics will be discussed separately in more depth.
Explicit instruction. The authors define the term explicit as, “direct, systematic, and comprehensive instruction to build phonemic awareness and phonetic decoding systems (phonics)” (p. 208). A study conducted by Torgesen and colleagues (1999) investigated the effects of interventions that varied in degree of explicitness on at-risk students’ phonemic awareness and word reading skills. The authors found that only the most explicit intervention produced the strongest and most reliable growth in word reading ability. The existing research evidence on explicit instruction is clear: as the instruction becomes more explicit, students’ word reading abilities increase (Brown & Felton, 1990; Foorman, & Torgesen, 2001; Iverson & Tunmer, 1994, Torgesen et al., 1999). In particular, students who enter school with weak phonological awareness benefit from explicit instruction (Juel & Minden-Cupp, 2000).

Intensive instruction. Next, it’s critical that at risk students receive intensive instruction. Students at risk for reading failure are not being afforded sufficient opportunities to respond to and benefit from instruction. In fact, some research suggests that up to 70% of the reading period is spent passively listening to the teacher and very little, if any, time is spent actively engaged in reading (Foorman, Goldenberg, Carlson, Saunders, & Pollard-Durodola, 2004). In order to become successful readers students must spend time in the classroom actively and intensively engaged in reading. Instructional intensity can be achieved two ways: increasing the total time spent in instruction or working with at risk children individually or in small groups. It is not always feasible for teachers to extend the learning period, nor is it necessary for all students. For students at the greatest risk of reading failure, small group instruction may be the best option since relatively little difference in outcomes scores is seen for students
in small groups versus students who receive one on one tutoring (Linan-Thompson & Hickman-Davis, 2002; Wise, Ring, & Olson, 1999). Although it may seem counterintuitive, it is not necessary to conduct intervention efforts in a strictly one-on-one format. The meta-analysis by the National Reading Panel (2000) showed that no additional gains were found with individual tutoring versus instruction in small groups of three to four students. Upon further reflection, however, this finding is plausible. Presumably, the reason that more intensified instruction promotes learning rates is not the topographical feature of the group (i.e., the number of group members), but is the frequency and rate of opportunities to respond to the instruction. There would be some, although very little, difference in opportunities to respond in a small group versus individual instruction. That difference in response frequency and rate would be negligible compared to the difference between whole class and small group. Students in small groups have many more opportunities to respond and the latency between responses is greatly reduced, therefore it is reasonable to expect that instruction in small groups results in higher learning rates, although no tangible additional benefit is gained from reducing a small group to individual instruction.

Torgesen and Foorman (2001) suggest several methods for increasing the intensity of instruction. Among the suggestions are ClassWide Peer Tutoring models (Greenwood, 1996), Peer Assisted Learning Strategies (PALS; Fuchs, Fuchs, Mathes, & Simmons, 1997), peer-mediated repeated readings (Staubitz, Cartledge, Yurick, & Lo, 2006) and other peer tutoring methods for developing basic early reading skills (Mathes, Torgesen, & Allor, 2001). All of these strategies offer flexibility of subject matter and increased opportunities to respond, making them highly adaptable and effective in
elementary classrooms. Furthermore, the use of peer-mediated strategies offers a simple method for increasing the intensity of instruction. A strategy designed to develop reading fluency and comprehension by intensifying reading practice is known as the Peer Assisted Learning Strategy (PALS; Fuchs et al., 2001). PALS is a peer-mediated reading activity that consists of three parts: partner reading, paragraph shrinking, and prediction relay. In partner reading students read to one another followed by a brief retelling. During paragraph shrinking and prediction relay gives students opportunities to summarize the main idea of the passage and make predictions. In a study with 33 first graders with and without disabilities in Title I and non-Title I schools, Fuchs et al. (2001b) examined the effects of the PALS plus an additional fluency program on reading performance. During intervention, the students were divided into three groups: a control group, PALS group, and PALS+Fluency. The intensity of instruction was greatest in the PALS+Fluency group and least intensive in the control group. The authors intensified instruction in both the PALS and PALS+Fluency group by increasing the opportunities to respond. However, instruction in the PALS+Fluency group was intensified even further by extending the total time in instruction. The classroom teachers modeled effective reading by demonstrating sound and word tasks including letter correspondence, blending and decoding, sounding out familiar words in stories, and sight word recognition to students in the PALS and PALS+Fluency groups. During the second half of the study, students in the PALS+Fluency group participated in an additional speed-reading game on the same tasks in three 1-minute time trials, while the other groups did not. The goal for the PALS+Fluency group was to increase their reading speed on each successive trial. After students in both the PALS and the PALS+Fluency groups were performing the sound and
word tasks independently, a partner reading activity was added. During partner reading, the students exchanged the roles of reading coach and tutee. Each partner read for 5 minutes during the 10-minute session. The students in the PALS+Fluency group made substantial gains, with .20 to .30 standard deviations above the other groups on reading fluency and comprehension after 2.5 hours of repeated reading tasks spread across a 22-week period. It is possible that this result can be attributed to the additional peer-mediated speed reading fluency building activity conducted with the sound and word tasks.

Supportive instruction. Finally, intervention efforts need to provide more supportive or scaffolded instructional delivery. Scaffolded instruction refers to instructional attributes that both cognitively and emotionally support the student in successfully completing an instructional task that he or she could not have otherwise achieved (Foorman & Torgesen, 2001). Instruction can be scaffolded two ways: (a) gradual skill-building in which students can demonstrate the prerequisite skills for each successively difficult task, and (b) a student-teacher dialogue in which the teacher guides the student to the successful completion of the task through specific and responsive questions and comments about the student’s behavior. Juel (1989) advocates this approach to supporting student learning because, “the ability to offer scaffolded support while children are acquiring reading skills may have increasing importance as the severity of the child’s disability increases” (Foorman & Torgesen, 2001, p. 210).

Effective Early Intervention Projects

When researchers and teachers employ the characteristics discussed above into their intervention strategies, students who struggle with reading can become effective and
efficient readers. Literature that supports direct, explicit, intensive, and supportive instruction at the individual school or classroom level is well represented in educational journals. Evidence of collaboration at multiple schools or school districts is limited. In order to make a significant impact on the epidemic of reading difficulties and disabilities, this type of effective early intervention needs to be adopted across schools and across districts. In recent years large scale studies have increased and although there are many quality studies worthy of discussion, it is beyond the scope of this review to cover them exhaustively. Two relevant large-scale intervention studies will be discussed here.

*Project PRIDE.* Using a three tiered, response to intervention model, Bursuck, Smith, Munk, & Damer (2004) implemented an early reading intervention in three urban elementary schools across Kindergarten and grades 1, 2, and 3. Project PRIDE is a comprehensive, data-driven, responsive intervention model that used systematic and explicit instruction on phonemic awareness, phonics, oral reading fluency, vocabulary, and comprehension to remediate and prevent further reading difficulties. At tier 1, students participated in whole class reading instruction using Harcourt Brace’s *Open Court* program (Adams et al., 2000). For students demonstrating mild difficulties, teachers made small, non-intrusive adaptations such as quickening the pace of instruction, modeling, guided practice, systematic errors correction, and an incentive system. As monthly or bi-monthly curriculum-based assessments indicated, some students entered tier 2, in which small group (2-8 students) was provided for 10 minutes per day in Kindergarten, or 30 minutes daily in grades 1 and 2. Letter sounds, phonemic segmenting and blending, decoding, sight word recognition, and fluency building was conducted according to the students’ needs and grade levels. Curriculum-based
assessments and DIBELS data were used to determine entry into tier 3, in which students participated for 30-50 minutes daily in *Reading Mastery*, a more systematic, supportive, and intensive intervention for reading difficulties. Classroom teachers and personnel conducted all interventions and were provided an on-site coach who modeled skills such as delivering praise statements, corrective feedback, and using appropriate pacing.

By the third year of the project (2003), there were 90 students who had participated in the project for all three years. At that time, PRIDE students were compared to comparable students from a control school on several measures. Significant differences were demonstrated in favor of the PRIDE students on both Nonsense Word Fluency (NWF) and Oral Reading Fluency (ORF). For NWF, 60.8% of PRIDE students were at benchmark, whereas only 32.9% of the control students attained this level of proficiency. On ORF, 31.1% of PRIDE students were at benchmark, compared to 17.8% of control students. Although these data indicate that many students were still at-risk in both groups, the control students presented a greater proportion of the students most at risk on both measures. Approximately 25% of control students were at risk on NWF and 63% were at-risk on ORF. In comparison, approximately 6% of PRIDE students remained at risk on NWF and 36% were at risk on ORF. It is clear that the PRIDE students made respectable gains in key early reading skills. However, the authors note that the structure of the model precluded students entering the most intensive third tier any earlier than February of the Kindergarten year. This was designed to provide every opportunity to succeed in the general education curriculum before intervening. Future research should investigate altering the model to provide access to the most intensive intervention at earlier stages of the project.
Taken together, several ambitious goals were achieved. First, many students who may have otherwise fallen further behind were provided instruction that assisted them in catching up with their peers on key reading skills. The fact that only 6% of PRIDE students remained at risk, while 25% of control students did is a strong endorsement for the program. Furthermore, by implementing a well-structured intervention project in three schools, the authors were able to reach more students at risk and still provide much-needed coaching and support for staff members. This is a crucial, but difficult task. Finally, the PRIDE model serves as an example of merging research with practice, an often elusive and intangible result (Fuchs & Fuchs, 1998).

*Project Optimize.* Scott Foresman’s Early Reading Intervention (ERI; Simmons, & Kame’enui, 2003) program was created as a byproduct of the intervention used in Project Optimize (Simmons et al., 2003). Ninety-six Kindergarten students who scored below the 20th percentile on letter knowledge and phonological awareness tasks were randomly assigned to one of three phonological awareness interventions. Participants in each group received 30 minutes of instruction daily for seven months. The most explicit and systematic intervention of the three was Optimize+Spelling, which focused on delivery of corrective feedback and maximized the number of opportunities for students to respond. Each of the Optimize+Spelling lessons was designed to incorporate the characteristics of effective intervention such as direct and explicit instruction, systematic sequencing, mediated scaffolding, as well as ensuring multiple response opportunities and immediate corrective feedback. The first 15 minutes of the Optimize+Spelling intervention was devoted to phonological awareness and alphabetic understanding activities identifying and discriminating sounds, blending, and segmenting. The second
15 minutes emphasized the connection between sounds, letters, and spelling. Activities included tracing and writing letters, writing the first, last, or middle sounds in words, and writing simple words (Consonant Vowel Consonant). The other two interventions incorporated much of the same content, but were not as systematic or explicit. *Open Court* (Harcourt Brace; Adams et al., 2000) was used in the study because it has been recognized as a promising research-based program for developing reading skills. The *Open Court* program included phonological awareness, handwriting, and decoding exercises as well as vocabulary practice and prediction activities to foster comprehension.

Students from the Optimize+Spelling intervention demonstrated the fastest gains in phonological awareness skills and alphabetic understanding. It is important to note that students from both Open Court and Optimize+Spelling groups demonstrated proficiency on the tasks; however, Optimize+Spelling students achieved proficiency about one month sooner than Open Court students. Furthermore, 48% of the most at-risk students achieved proficiency in alphabetic understanding in the Optimize+Spelling program, whereas 13% of the most at-risk students in Open Court achieved this goal. It is also worth noting that the differences between the groups on several measures such as word attack, word identification, letter naming, spelling, and oral reading fluency were found to be statistically significant. Finally, one Reading First coordinator from a school that participated in the project noted that although the school’s demographics had not changed in three years, referrals to special education had dropped from 14% to 3-4% (Bell, 2003).

*Early Reading Intervention and Prevention of Secondary Disabilities*

The need for early reading intervention and prevention of reading difficulties is crucial since students who show early deficits in reading are at risk for secondary, or
comorbid disorders, particularly behavior disorders (BD; Coie & Jacobs, 1993; Walker, Colvin, & Ramsey, 1995; Lane, 1999; Lane et al., 2002). Several researchers have proposed directional hypotheses regarding the causal relationship between behavior problems and academic underachievement (Hinshaw, 1992; Bower, 1995). Hinshaw proposed three models:

1. Poor achievement, particularly in reading, leads to externalizing problem behaviors. This model implies that when students are faced with task demands that are too difficult, they will engage in disruptive behavior in order to escape the aversive condition. In this case, academic intervention would be an appropriate response to the situation.

2. Problem behaviors lead to academic underachievement. Since many problem behaviors are incompatible with academic behaviors (e.g., throwing a pencil vs. writing spelling words), it is likely that a disruptive student would miss a great deal of pertinent content and therefore perform poorly on academic tasks. In this case, behavioral intervention would be the most appropriate course of action.

3. Both problem behaviors and underachievement patterns are continually influencing one another. The reciprocal nature of this model indicates that causal links are not solely unidirectional and fixed. Both underachievement and problem behavior can be responsible for the student’s poor outcomes and either or both could be the salient cause at any given time. The intervention action in this model should address both behavioral deficits and academic deficits.
Regardless of the causal model, it is clear that both behavioral and academic problems lead to poor outcomes for these students. Directing the emphasis for at risk readers to intensive reading instruction is critical since there is a strong correlation between reading achievement and externalizing behavior patterns. As a result, the remediation of one area (reading) may have an effect on the other area, in this case, behavior (Lane, Wehby, Menzies, Gregg, Doukas, & Munton, 2002). This approach is a promising one since many studies of academic deficits and behavioral problems indicate that remediating academic deficits may produce collateral effects on behavior (Coie & Krehbiel, 1984; Lane, O’Shaughnessy, Lambros, Gresham, & Beebe-Frankenberger, 2001). The case for early reading intervention is strengthened by the notion that not only can academic deficits be resolved, thus potentially preventing future behavior problems, but students who already demonstrate behavioral issues may essentially receive twice the benefit for one intervention.

A review of the recent literature (1990-2003) by Lane (2004) on academic interventions for students with BD yielded surprisingly few published controlled studies. A total of 25 articles met the following inclusion criteria:

1. The intervention must be school-based and directed by an interventionist, teacher, or peer.
2. School-age children with or at risk for BD were the participants.
3. The article was published between 1990 and 2003.
4. The article reported outcomes that were academic performance or academic-related behaviors.
However, of these 25 articles, only 14 were related to reading interventions. The others were mathematics or writing interventions. Three of these 14 studies focused specifically early reading behaviors such as phonological and phonemic awareness skills in small group, classroom-based interventions delivered by a graduate student or school personnel. The three studies will be discussed separately in chronological order.

*K.L. Lane (1999).* The effects of an academic intervention and a social skills intervention were compared for improvements in academic performance, social competence, and reductions of problem behavior. A total of 53 first-grade students at risk for problem behavior participated in the study. Groups were randomly assigned as intact classrooms to avoid classroom disturbances to the extent possible. Six classrooms from two separate school sites were randomly assigned to one of these conditions: academic intervention, social skills intervention, or the treatment contact control group. Each group met four days per week for 30 minutes, resulting in a total of 12.5 hours of training.

The intervention used in the academic condition was the Phonological Awareness Training for Reading Program (Torgesen & Bryant, 1994). This scripted program consists of four elements: warm-up, sound blending, sound segmenting, and reading and spelling. The social skills intervention group received scripted instruction from the *Social Skills Intervention Guide: Practical Strategies for Social Skills Training* (Elliot & Gresham, 1991). Lesson formats includes five phases: Tell, Show, Do, Follow Through and Practice, and Generalization. Lessons for each group were matched to the area of identified social needs. The treatment contact control group received general academic assistance in areas such as jobs in the community, the five senses, and measures of time. The purpose of this group was to control for the effects of time and attention.
The outcome measures were conducted at three points during the study: prior to onset of intervention, at the conclusion of the intervention, and three weeks post intervention as a follow up measure. Behavioral and social skills outcomes were measured by the SSRS and the Critical Events Index (CEI; Walker & Severson, 1992). Academic skills were measured by the Woodcock Reading Mastery Test – Revised (Woodcock, 1987), The Test of Phonological Awareness (TOPA; Torgesen & Bryant, 1994), and The SSRS – Academic Competence (Gresham & Elliot, 1990).

A mixed model, repeated measures design was used with students nested within classrooms and classrooms (n=6) nested within treatment conditions. A repeated measure Analysis of Variance (ANOVA) was conducted and if the results were significant a post hoc analysis was done. To control for group differences at the outset, an Analysis of Covariance (ANCOVA) with the pretest as the covariate was conducted. Finally, a series of one-way ANOVAS were conducted on each dependent measure for the post-intervention and follow up point, in order to determine if there were differences in maintenance between these two points in time for each group. (For a complete report of results see Lane, 1999).

The results did not support the hypothesis that academic intervention affects both academic and social skills. There were no significant differences found between the intervention groups, indicating that these groups did not change at significantly different rates from one another. However, significant main effects were found for time on the reading measures, indicating that the six-week intervention effected some change on all participants’ reading. It is possible that the changes were a function of time. The results of the ANCOVA showed a significant improvement for the academic skills group on the
TOPA. This result suggests that the academic skills group made gains on phonological awareness, although similar gains were not seen for the word attack and social competence outcome measures. Finally, no significant changes were found on any dependent measures between post-intervention and follow-up. This indicates that whatever changes did take place were maintained over the three week post-intervention period.

Although there were some gains on the phonological awareness component for the academic skills group, many other potential gains were not detected by the statistical measures employed for this design. The small number of students in each treatment condition is a considerable limitation when using statistical analyses and surely contributed to a lack of power in identifying statistically significant changes. Related to this limitation is the hierarchical nature of the design. Since the classroom (rather than the individual) was the unit of analysis, the power of the tests suffered additional weakening. It is understandable that the classroom was used as the unit of analysis and nested within conditions in order to prevent major disruptions for the teachers, but certainly this design compromised the power of the statistical tools used. It is also worth noting that the SSRS is not designed to be a pre-post measure in such a short span of time. The author suggested that more sensitive measures, such as direct observation, would help identify smaller, but important, behavior changes, which could then lend itself to a statistical analysis.

K.L. Lane, T.E. O’Shaughnessy, K.M. Lambros, F.M. Gresham, & M.E. Beebe-Frankenberger (2001). Although having or being at risk for BD can potentially put students at risk for academic difficulties, Lane et al. caution that academic deficits,
particularly in reading, are even more likely for students who demonstrate comorbid conduct disorders and inattention. This population has substantial problems with phonological awareness, decoding, and reading comprehension. In light of this risk, the purpose of this research study was to examine the efficacy of a phonological awareness training program for first grade students at risk for conduct disorders and inattention problems as well as reading failure. In order to be included in the study, three criteria needed to be met:

1. Student scored below the 25th percentile on the TOPA (Torgesen & Bryant, 1994).

2. An SSRS – T (Elliot & Gresham, 1991) score at or above the 75th percentile or A Critical Events Index (CEI; Walker & Severson, 1992) of 1 or more.

3. The SSRS – T score Externalizing and Hyperactivity subscale scores exceeded gender norms by one standard deviation.

Seven children met all three criteria and were subsumed by a study that included an additional 13 children who met the first two criteria. The participants were from two different sites: Arizona and Georgia. Six students each were assigned to one of two intervention groups in Arizona and four students each was assigned to two intervention groups in Georgia. Each group received intervention for 30 minutes a day, three days per week for a total of 10 weeks. This resulted in 15 hours of training total.

In addition to the SSRS – T, CEI, and the TOPA, direct measures of reading and social behavior were collected during the study. Data were collected prior to the onset of intervention, directly after the intervention concluded, and at a two-week follow-up point. Weekly probes were also collected during intervention. The Nonsense Word Fluency
The (NWF) subtest of the Dynamic Indicators of Basic Literacy Skills (DIBELS) was used to assess the students’ word attack skills. The direct observation measures of classroom behavior assessed (a) Total Disruptive Behavior (TDB) and (b) Negative Social interactions. TDB included any behaviors that disrupted the classroom ecology such as hitting, biting, choking, being out of seat, touching or grabbing other students. Negative social interactions included quarrelsome behavior, taunting, aggressive threats, name-calling, or bossy commands. The experimental design was a multiple baseline across intervention groups. Three of the seven children who met all three inclusion criteria comprised the first intervention group and the remaining four students were in the second intervention group. In addition to visual analysis of the data, a mean scores comparison across phases and effect sizes were calculated.

The results of the mean changes by phase comparison yielded substantial gains for all students in word attack skills as measured by the DIBELS NWF subtest and these gains were maintained. Visual inspection of the TDB data indicated that all students demonstrated a reduction in problem behavior from baseline to intervention. However, only one continued to show decreases in TDB at the follow-up point. For at least three of the students, a clear reciprocal relationship was demonstrated between word attack skills and TDB. All students, except one, showed decreases of NSI, but only two demonstrated continued improvement in the post-intervention and follow-up phases. Effect sizes for the DIBELS NWF ranged from 1.22 to 3.81 with the three oldest students demonstrating the most robust improvement with effect sizes ranging between 2.75 and 3.81. Interestingly, the same three oldest students demonstrated the least improvement on TDB with effect sizes ranging between -.01 and -.18.
The overall results of this study support the use of an early reading intervention for students who are at risk for reading failure and comorbid conduct disorders and inattention. Although somewhat less substantial than the improvements for word attack skills, some general improvement was demonstrated in classroom behavior as well. However, the authors caution that the intervention may not have been of sufficient length and intensity to effect lasting changes on TDB. Additional behavior management support would have been appropriate. It appears that maladaptive behaviors, especially in older children, are more resistant to intervention and need to be remediated early. Findings from the visual inspection and the effect sizes support a preliminary conclusion that a reciprocal relationship does exist between improvements in phonemic awareness skills and reductions in maladaptive behaviors.

*K.L. Lane, J.H. Wehby, H.M. Menzies, R.M. Gregg, G.L. Doukas, & S.M. Munton (2002).* Lane’s previous research examined the efficacy of a phonological awareness training program for first grade students at risk for BD and reading failure. Her findings point to a negative correlation between phonemic awareness and disruptive behavior. However, in each of the previous studies, a research associate from the university served as the interventionist. Although using university personnel in the role of interventionist may yield strong treatment integrity data, the social validity of such findings in an applied sense remains unknown since school-based personnel were not directly involved with the intervention delivery. Of course, the most desirable intervention model would include school-based personnel as the interventionists, with university assistants serving as coaches. The authors extended their previous research by using the school literacy leader, who was also a doctoral candidate in special education, as the interventionist.
Students who were nonresponsive to a school-wide literacy and behavioral intervention program were considered for inclusion in the study. Additionally, participants had to meet the following criteria:

1. Student received an at risk rating for antisocial behavior according to the SSRS – T form.
2. Following three months of the school-wide literacy and behavioral intervention, the student’s literacy skills were in the bottom 33% of the class.

A total of seven students met the inclusion criteria and were randomly assigned to one of two intervention groups. Each group met for 30 minutes of instruction three to four times per week for nine weeks. This yielded a total of 15 hours of intervention. The intervention program was John Shefelbine’s Phonics Chapter Books (Shefelbine, 1998), which provided explicit instruction on phonemic awareness, sound-syllable connections and high frequency sight words, chapter book reading, and writing and dictation.

A number of outcome measures were collected. Similar to previous studies by Lane, standardized measures such as the TOPA (Torgesen & Bryant, 1994), SSRS (Gresham & Elliot, 1990), and the CEI (Walker & Severson, 1992) were collected. Additionally, the nonsense word fluency (NWF) and oral reading fluency (ORF) subtests from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Kaminski & Good, 1996) were employed as outcome measures. As in the previous study, direct observations of Total Disruptive Behavior (TDB) and Negative Social Interactions (NSI) were conducted.
The experimental design was a single subject multiple baseline across intervention groups. NWF, ORF, TDB, and NSI were assessed for each participant during baseline, intervention, post-intervention, and follow-up. The data were analyzed by visual inspection, mean score comparisons across phases, and effect size calculations. It should be noted that two students who began the study did not complete it. Therefore, the authors chose to adjust for this attrition by examining treatment outcomes of one student from each group who demonstrated similar academic and behavioral profiles at the beginning of the intervention. Effect size calculations were conducted at the individual level, rather than for the group.

Results of this study indicate that all participants made strong progress on word attack skills. Effect sizes for NWF ranged from .59 to 3.61. Effect size results for ORF were mixed. Three participants demonstrated gains, but four actually had moderate decreases. Yet, the mean score comparison between baseline and intervention showed an increase in ORF for each participant. The effect sizes for the behavioral measures showed a marked reduction in disruptive behavior in the classroom. The effect size range for TDB was between –0.58 and –1.58. Contrary to findings from the previous study, the post-intervention mean for all 5 students on the TDB measure was zero. Furthermore, all but one student had decreases in NSI. The effect sizes for this outcome measure ranged from –1.66 to 0.67. When graphs of TDB and NWF were juxtaposed for two of the students, a clear inverse relationship can be seen in the data. As NWF increased, TDB decreased.
The results of this study confirm the findings of Lane’s previous work by demonstrating an inverse relationship between phonemic awareness and maladaptive behavior. This study extended the line of inquiry by employing school-based personnel to implement the intervention, rather than using university personnel as interventionists. Although the classroom teacher was not the lead interventionist, social validity data were collected from her regarding her views on the treatment’s effectiveness and utility. The classroom teacher rated the procedures as acceptable, but did not think the intervention was appropriate to address academic and behavioral problems. Moreover, she claimed that the intervention did not effect any substantial change on her students’ behavior. However, this claim ran in direct opposition to the results on the outcome measures and direct observations. In fact, most students’ disruptive behavior was high and variable in baseline and decreased to zero at the conclusion of the intervention. Nevertheless, the results of this study support the conclusion that phonemic awareness and phonics training can substantially improve early literacy skills and concomitantly decrease maladaptive behaviors, thus serving as a protective factor against both behavioral and academic failure.

A more recent study (Nelson, Stage, Epstein, & Pierce, 2005) examined the effectiveness of an early, intensive prereading intervention on social and reading behavior change. A total of 63 Kindergarten children at risk for BD and reading deficits from 27 different classrooms and 10 different schools participated in the study. Participants were selected using a three-step procedure:
1. Classroom teachers generated two rank-ordered, mutually exclusive lists of children who demonstrated patterns of internalizing or externalizing problem behaviors. 

2. Teachers completed three scales of the Early Screening Project (ESP; Walker, Severson, & Feil, 1995) for the five highest-ranking children on both the externalizing and internalizing behaviors lists. The scales were: (a) Critical Events Index, (b) Maladaptive Behavior, and (c) Adaptive Behavior. 

3. The Letter-naming fluency portion of the DIBELS was administered to children meeting the normative criteria for BD. Students who scored 60 or more on the ESP and named fewer than seven letters on the DIBELS probe were eligible for the study. 

A pretest posttest group design was used to evaluate the effects of the Stepping Stones to Literacy (Nelson, Cooper, & Gonzalez, 2004) on the reading and social behaviors of the participants. The Stepping Stones to Literacy program contains 25 lessons on the following skills: (a) identification, manipulation, and memory of environmental sounds; (b) letter names; (c) sentence meanings; (d) phonological awareness; (e) phonemic awareness; and (f) serial processing and rapid automatic naming tasks. Each lesson lasted between 10 and 20 minutes and was delivered daily. All participants in the study continued to receive their regular daily reading instruction in the classroom. The nonspecific treatment condition did not receive any additional or alternate intervention.
Four literacy and three social behaviors were the dependent variables. The literacy behaviors included: (a) Phonological awareness (PA) as measured by the Comprehensive Test of Phonological Processing (CTOPP, Wagner, Torgesen, & Rashotte, 1999); (b) Word reading as measured by the Word Identification (WI) and Word Attack (WA) subtests of the Woodcock Reading Mastery Tests – Revised (Woodcock, 1987); (c) Letter naming fluency (LNF) as measure by the corresponding subtest of the DIBELS; and (d) Rapid automatic naming as measured by the Rapid Naming (RN) subtest of the CTOPP. The three social behaviors measured were: (a) School functioning (SF), (b) Interpersonal strength (IR), and (c) Intrapersonal strength (IS). All social behavior subtests came from the Behavioral and Emotional Rating Scale (BERS, Epstein & Sharma, 1998).

Children in the treatment condition demonstrated overall improvement in all areas of literacy measured. Children in the nonspecific treatment group showed little or no improvement. Significant main effects for change scores (Group X Change) were demonstrated in all but one area (CTOPP RN). Neither group demonstrated significant changes in Rapid Naming skills. The obtained effect sizes for PA, WI, WA, RN, and LNF skills were .55, .99, .92, .07, and .79, respectively. On the other hand, no statistically significant changes for Group X Change interactions were obtained on any of the BERS subtests. The effect sizes for SF, IR, and IS were –.56, 0.0, and -.69, respectively. These findings suggest that an intensive, prereading intervention program for students that show reading and behavioral risk can substantially improve early reading skills, but not social behaviors. This is noteworthy because students with or at risk for BD have often been resistant to early literacy interventions (Al Otaiba & Fuchs, 2002).
A longitudinal study conducted by Gunn and colleagues (2005) represents an interesting intersection of several topics related to this dissertation including: (a) the use of an explicit, intensive, and supportive early reading intervention for children at risk for reading problems, (b) the collateral effects of reading intervention on prevention of behavior problems and disorders, and (c) the extent to which reading interventions that adhere to the principles of effective instruction can be effective for ELLs and non-ELLs.

The purpose of the study was to examine the effectiveness of a supplemental early reading intervention on several reading behaviors over a four-year period in four Oregon communities. Due to the frequent coexistence of reading deficits and problem behaviors, the authors purposively indicated an inclusion criterion of aggressive behaviors. Additionally, since Latino students are increasingly represented in classrooms and experience academic failure at high rates, this population was specifically targeted for analysis. Students were assessed at the beginning of Kindergarten and participants were assessed in the spring for four years. Reading skills were evaluated using the Woodcock Johnson Tests of Achievement (Woodcock & Mathur, 1990) Word Attack and Word Identification subtests as well as an oral reading fluency test. Behavior was assessed using the Walker-McConnell Test of Social Skills (Walker & McConnell, 1988).
The resulting sample included 148 students; 80 students demonstrated risk on reading skills only and 68 students were at risk for reading and behavior problems. Nine instructional assistants (IAs) received 10 hours of training on delivering the instruction and managing and motivating student behavior. Additionally, the authors presented an overview of reading research and effective instruction for reading acquisition. Three of the IAs were certified teachers and seven had previous experience tutoring elementary students or working in small group instructional settings.

The interventions used were *Reading Mastery* (Englemann & Bruner, 1988) for participants in Kindergarten, first, or second grade, or *Corrective Reading* (Englemann, Carnine, & Johnson, 1988), for participants in third or fourth grade. Students received the interventions for 50 minutes three days per week. The intervention time was divided among phonics, word reading, spelling, and repeated readings to build oral reading fluency. If necessary, additional time was spent with Latino students who needed extra instruction with English vocabulary. Improvements in all three measures were demonstrated for treatment students. Two years following intervention, differences were still evident by condition for Word Identification, Word Attack, and oral reading fluency. Interestingly, even though Latino students had lower baseline scores, they benefited as much or more than their native English-speaking peers. The authors also note that their entry level of English proficiency had no effect on their ability to benefit from the intervention and that schools should not delay instruction until these students develop English oral language, but rather introduce instruction early. There were no significant difference found between students selected for reading deficits only and students selected for reading deficits and behavior problems. However, the students with reading and
behavior problems showed greater gains than their matched controls, suggesting that students with both reading and behavior problems can benefit from supplemental instruction. Finally, the authors concluded that no significant differences existed for grade and intervention. They found that both groups, Kindergarten through second grade Reading Mastery students and fourth and fifth grade Corrective Reading students, benefited similarly.

The line of research just described is among the most cohesive and current work in the area of academic and behavioral interventions for students with or at risk for BD. Even though the quantity of research in this area is severely limited, a few key points can be drawn from the work described above:

1. It is possible that early reading interventions can have some positive collateral effects on maladaptive behavior, but the results of research into this relationship so far have been inconclusive (Barton-Atwood, Wehby, & Falk, 2005). The reciprocal relationship between early literacy skills and behavior was not as clear and robust as might have been preferred in both the Lane et al. (1999) and Nelson et al. (2005) studies. However, as previously stated, the small sample size restricted the power of the statistical tests used to draw conclusions. The author justifiably noted that changes in behavior did in fact occur and that those changes would be more visible through direct observation data. In Lane et al. (2001; 2002) the direct observation data did verify that behavioral changes were taking place concomitantly with academic intervention. However, the magnitude of behavioral change as a result of academic intervention may not be sufficient to quell the most problematic
classroom behaviors. Academic intervention has been shown to positively affect maladaptive behavior, but may not be powerful enough to reduce it to minimal levels. Some additional behavioral intervention may be necessary for a small portion of students, but considering that this work was done for students at risk for academic and behavioral problems, an academic intervention approach is very appropriate.

2. Supplemental reading instruction was shown to be an effective tool for students at risk for reading and behavior problems. The interventions were relatively low cost, time efficient, and unobtrusive. In fact, the intervention period in the Lane studies amounted to a total of only 15 hours. Although the students did make gains in such a short period and the classroom teacher’s duties were never imposed upon, the applicability and utility of such a model is questionable since the intervention work was mostly performed by university personnel. That is, would the procedures fit seamlessly into the regular classroom routine? Transference of the interventionist responsibility to school personnel is the ultimate goal of most school-based research so that the benefits of the program can become a permanent and continuous part of that school. Lane et al. (2002) addressed the personnel issue by using the school literacy leader as the interventionist. This adjustment brought the research one step closer to being adopted into the school; however, the classroom teacher still expressed some dissatisfaction with the program. Interestingly, the teacher’s perceptions of student performance actually ran contrary to the data on some occasions. This finding illuminates the need to not just use school
personnel as interventionists, but also enlist them as stakeholders in the programs so that teachers themselves are not only aware of actual progress in student performance, but also maintain some sense of personal responsibility for it. It is not enough to use auxiliary members of the school faculty. The teachers must become key players in the intervention work by implementing and observing the programs and, more importantly, actively monitoring student response to and progress in that program.

3. The study by Gunn and colleagues provides convincing evidence supporting the use of early reading interventions as a tool for the prevention of both reading problems and behavior disorders. Furthermore, this study supports the contention that early reading interventions in English can be as effective, if not more so, for ELLs. It appears that ELLs and non-ELLs can both benefit from similar types of intervention.

Conclusion

Supplemental reading instruction for children at risk for reading failure improves skill level of struggling readers and may prevent future problems in reading and other academic skill areas (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Snow, Burns, & Griffin, 1998; Torgesen, Wagner, & Rashotte, 1997b; Vellutino et al., 1996). Furthermore, the literature suggests that some relationship exists between early reading intervention and the prevention of troublesome behaviors, which may complicate or exacerbate achievement problems. Though Hinshaw’s model (1992) provides three different directional relationships, the literature discussed above, as well as perspectives of leading authorities, lend support to the advantages of early reading intervention over
solely behavioral ones (Kellam, Mayer, Rebok, & Hawkins, 1998) and these findings appear to be applicable to ELLs (Gunn, Smolkowski, Biglan, Black, & Blair, 2005).

Reading Acquisition for English Language Learners

Over the past 20 years, the representation of students who are English Language Learners (ELLs) has increased dramatically and their numbers continue to increase at an accelerated pace (Gunn, Smolkowski, Biglan, Black, & Blair, 2005; Klingner, Artiles, & Barletta, 2006; Snow, Burns, & Griffin, 1998). ELLs account for approximately 6% of the school age population and the representation of Spanish-speaking students among that population has been estimated between 70% and 80% (Fitzgerald, 1995; Gunn, Smolkowski, Biglan, Black, & Blair, 2005; Haagar, & Windmueller, 2001; Klingner, Artiles, & Barletta, 2006). Some estimates indicate that as many as 43% of teachers reported having at least one ELL in their classroom in 2002 (NICHD, 2003). It is clear that the nation’s students are becoming increasingly diverse and schools need to become more responsive to these changing demographics (Wilkinson, Ortiz, Robertson, & Kushner, 2006).

Students who are learning English as a second language or report another language as the primary language in the home, or ELLs, present special academic risks including underachievement, grade retention, and attrition from school (Abedi, 2002; August & Hakuta, 1997). In the early grades, reading acquisition is particularly challenging for these students (Haagar, & Mueller, 2001). A special report from Zehler and colleagues (2003) on the achievement status of limited English proficient students suggests that nearly three quarters of ELLs read below grade level in English at the third grade and more than half perform below grade level in math. A typical response to this
type of underachievement or lack of response to classroom instruction is referral to special education. In fact, approximately 56% of ELLs being served in special education are referred for reading problems and 24% are served for a speech or language impairment (NICHD, 2003). Furthermore, the rate of placement in special education appears to be negatively correlated with the level of English proficiency. That is, as English proficiency increases, the rate of placement in special education decreases. However, we know from research on non-ELL populations that not all children with reading difficulties have learning disabilities in reading or any other areas (McCardle, Mele-McCarthy, & Leos, 2005). Finally, ELLs with disabilities are more likely to be instructed in more restrictive settings, receive fewer language supports, and are seriously overrepresented from fifth grade and throughout the secondary level, indicating more long-term placements and less movement out of special education (Artiles, Rueda, Salazar, & Higareda, 2005; Klingner, Artiles, & Barletta, 2006).

Although the literature is fairly consistent on early reading intervention for native English speakers, there is much less agreement about reading intervention for ELLs. Some researchers have suggested that literacy instruction in a second language should occur only after literacy is established in the primary language, while others contend that putting off secondary language instruction is an unnecessary delay (Fitzgerald, & Noblit, 1999; Lenters, 2004). Regardless of the timing of instruction, there is one area of clear convergence emerging from research. As in early reading intervention for non-ELLS, phonological awareness and the ability to detect phonemes (i.e. phonemic awareness) in the primary language is a strong predictor of subsequent performance on parallel reading and language tasks in English (Cisero & Royer, 1995). Furthermore, studies suggest that
phonological awareness can be transferred from one’s first language to a second (Chiappe & Siegel, 1999; Durgonoglu, Nagy, & Hancin-Bhatt, 1993). Still, many questions exist regarding the most beneficial route to literacy in both the primary language and English for ELLs. Should ELLs who demonstrate reading difficulties receive intervention in their primary language if the core instruction in the classroom is in English? Or is it more important that the core instruction and intervention strategies maintain language congruence, whether in the primary language or English?

Advocates of primary-language instruction contend that students who learn literacy in their native language first can make a better transition to English at a later time (Thomas & Collier, 1997), while those who support primary instruction in English only claim that providing intensive, supportive, and continuous instruction leads to a higher proficiency in oral and written English (Rossell & Baker, 1996). Indeed, the evidence is mixed. Studies supporting both viewpoints exist, and the most current direction in the ELL literature appears to be a bilingual approach, in which the goal is literacy in both the primary language and English (Foorman, Goldenberg, Carlson, Saunders, & Pollard-Durodola, 2004). Methodological questions regarding bilingual instruction include the length and amount of primary language instruction needed, the logistics of staffing enough bilingual professionals in schools, and congruence of supplemental intervention programs with the primary literacy instruction. The studies reviewed in this section will examine the effectiveness of two different intervention models (e.g. English-language or primary-language) on phonological awareness and other early reading skills in both English and primary languages.
English Language Reading Intervention

Since reading is a developmental process that requires different emphases at different stages, the amount of support in the primary language may vary as the student becomes more literate. Some argue that since beginning reading requires more skill-building of phonological processes, rather than context or meaning-based processes, ELLs can successfully navigate these activities with little detriment due to minimal English vocabulary (Foorman et al., 2004). Therefore, core reading or supplemental instruction in phonological awareness may be conducted in English.

In a multiple-probe across participants design, Jitendra and colleagues (2004) investigated the effects of the Read Well (Sprick et al., 1998-2000) program on seven students’ alphabetic understanding skills such as phonemic segmentation fluency, letter naming fluency, letter sound fluency, nonsense word fluency, oral reading fluency, and word reading skills measured by the Woodcock Johnson – Revised (1987) Word Attack and Letter Word Identification subtests. The Read Well program is designed for beginning or remedial readers whose instructional reading level is at the first grade. There are 38 units that can be taught in 2-, 3-, 4-, or 6-day cycles, according to the student’s learning pace. The instructional strategies are based on effective reading instruction research and employ systematic, explicit, and direct instruction on phonics, phonemic awareness, oral reading fluency, vocabulary, and comprehension. The decoding skills covered by Read Well include segmenting, blending, rhyming, sight word acquisition, and fluency building, with multiple opportunities to practice new skills. At the end of each unit students are assessed and do not move on until mastery on the unit’s skills is demonstrated. Teachers participated in a 2-hour training sessions in which lessons were
modeled and practiced followed by feedback from the trainers. The teachers also viewed a videotape of a *Read Well* lesson and were provided an opportunity to read and practice lessons before the intervention began.

Results for the six non-ELL and one ELL participants generally suggest that *Read Well* is an effective intervention for improving early reading skills. All students improved on LSF tasks, and most students (80%) improved their NWF, but only 60% improved on LNF and ORF. It is interesting to note however, that the ELL student demonstrated the most robust effect size for ORF, 1.5, whereas the others had relatively diminished, but still moderate effect sizes. The results of the Woodcock Johnson subtests produced mixed results in which only 40% of students showed improvement. The authors concluded that the intervention was generally effective, and specifically beneficial for the ELL student, and posited that the minimal length of intervention (between 2 and 7 weeks) may not have been of sufficient length to remediate reading deficits for the “treatment resisters”.

Another study (Denton, Anthony, Parker, & Hasbrouk, 2004) examined the effectiveness of both the *Read Well* and *Read Naturally* (Ihnot, 1992) programs for a sample of 93 Spanish-English bilingual students. Participants had to have adequate oral English proficiency and a basic proficiency in reading Spanish. The sample was drawn from 17 bilingual classrooms and divided into four groups: *Read Well* intervention group, *Read Well* comparison group, *Read Naturally* intervention group, and *Read Naturally* comparison group. The sample included 22 students in second grade, 37 in third grade, 28 in fourth grade, and six in fifth grade. Because the students spanned a range of ages and grades, they were divided into two groups, emergent or established decoding ability, based on results from the Word Attack subtest of the Woodcock Johnson Reading
Mastery Tests – Revised (WRMT-R, Woodcock, 1987). The participants’ schools were part of a transitional bilingual program, so the language of primary instruction in the classroom varied. Some teachers reported that instruction was in English only, while other teachers reported that instruction was only in Spanish, and others stated that instruction was in both Spanish and English.

The students received either Read Well (intervention previously described; Jitendra et al., 2004) or Read Naturally interventions for 40 minutes three times per week for 10 weeks. Sixteen students were in the Read Well treatment group (14 comparisons), and 29 participated in Read Naturally (25 comparisons). Read Naturally contains fluency building activities such as repeated readings, vocabulary and comprehension instruction, and progress monitoring. Students typically read a passage of 75-175 words three times with the assistance of an audiotape, followed by repeated readings without the tape. Attention to comprehension is emphasized by written retelling or multiple choice comprehension questions. The authors modified the vocabulary activities by having tutors select two high frequency words from the passages and teach them to the students. Decoding was addressed by previewing the passage and teaching unknown words to the student. Results for the Read Well program indicated that there was a significant interaction between time and group on the Word Identification subtest only. Furthermore, 16% of the variance was attributable to the program. Although, not statistically significant, 6% of the variance on the Word Attack subtest was accounted for by group assignment. No statistical significance and minimal effect sizes were found for group by time interactions for all variables. Overall, the authors concluded that the Read Well program was successful in improving English word reading ability for bilingual students,
and that their findings did not support the prediction that the Read Naturally program would result in improved English decoding skills for tutored students. The authors recommend that teachers help students realize how English differs in orthography and phonology from the students’ primary language and adapt this knowledge to early attempts at English reading. Teachers should not assume that reading skills in the primary language will automatically transfer to English.

*Primary Language Intervention*

Some researchers contend that early reading intervention for ELLs is most effective when the language of the intervention is matched to the language in the classroom and is the student’s primary language. Furthermore, primary-language advocates believe that after developing expressive and receptive language and early reading skills in the primary language, these skills can best be transferred to English more easily.

To illustrate this point, Vaughn and colleagues (2006) conducted a study with 64 (31 treatment, 34 comparison) first grade ELL students from seven different schools. The authors evaluated the effectiveness of a supplemental intervention that matched the students’ language of core instruction (both in Spanish) on Spanish reading skills and oral language skills in both Spanish and English. Participants were randomly assigned to either treatment of comparison groups. The comparison group received the school’s standard intervention for poor readers. Outcome measures included the Letter-Word Identification subtest from the Woodcock Language Proficiency Battery – Revised (Spanish form; Woodcock & Munoz-Sandoval, 1995), the Spanish version of the DIBELS (*Indicadores Dinámicos del Éxito en la Lectura; IDEL*; Good, Bank, & Watson,
2003), and the *Test of Phonological Processing*, Spanish version (TOPP-S; derived from the *Comprehensive Test of Phonological Processing; CTOPP*; Wagner, Torgesen, & Rashotte, 1999). Students received 50 minutes of intervention three to five times per week with a certified bilingual instructor. The instructors participated in 12 hours of training prior to the onset of the study, six hours of retraining, and an additional one to two hours of staff development on site. On site coaching varied from weekly to monthly depending on need.

The intervention design followed the principles of effective instruction of non-ELL students. Lessons were delivered at a fast pace with multiple opportunities to respond and receive instructional feedback. Rather than a focus on phonemes, the material focused on syllabic structure, as is reflective of the structure of the Spanish language. Results indicated that intervention students made gains on nearly all measures and significantly outperformed comparison students on Spanish measures of word attack, phonemic awareness, word reading, comprehension, fluency, and oral language proficiency. Furthermore, comprehension gains extended to written Spanish material, which may be accounted for by the transparency of Spanish orthography (i.e. fewer irregularities) or the concomitant gains in fluency. Also, both treatment and comparison students demonstrated growth on English oral language and reading outcomes, suggesting that Spanish reading skills may transfer to or support language and reading development in English. The authors concluded that the improvements in so many skills may be attributable to the comprehensive approach to instruction, including oral language, vocabulary, decoding, fluency, and comprehension activities.
Considerations for ELL reading interventions

One important consideration for providing early reading interventions for ELLs is the linguistic diversity found in many school’s bilingual or ELL programs. Often, these classrooms are in urban settings and serve students from many different linguistic and cultural backgrounds. Given that resources and funding are already scarce in many urban districts, it is unlikely that professionals from every represented language background can be hired to serve these students. Therefore, more research is needed to determine if ELLs who need early reading intervention can be instructed in English.

Lesaux and Siegel (2003) conducted a longitudinal study that followed 978 students from Kindergarten through second grade in a Canadian school district of 30 schools. The participants included 188 ELL and 790 non-ELL students. The participants in this study differ from other studies on ELL reading achievement in that several linguistic backgrounds including Persian, Mandarin, Korean, Polish, Cantonese, and Farsi are represented. Participants who scored at or below the 25th percentile on the Wide Range Achievement Test 3 (WRAT3; Wilkinson, 1993) were considered at risk for reading failure. Two hundred and ninety-six students (236 Non-ELL and 60 ELL) qualified as at risk. Students were again tested at the end of second grade and results indicated that only seven ELL students were classified at reading disabled whereas 33 non-ELL students were reading disabled.

A classroom-based small group phonological awareness instructional program was provided for students at risk in Kindergarten. Small groups included both ELL and non-ELL students based on comparable phonological awareness abilities. The intervention was provided 3-4 days a week for 20 minutes. The regular classroom reading
program was described as a “balanced literacy” approach, which embedded phonological awareness activities in the curriculum. A phonics program was in place for students in first grade that continued to demonstrate risk. The data for ELL and non-ELL students were analyzed separately using hierarchical regression methods. Several predictor variables were regressed in a step-wise fashion on the second grade WRAT3 reading score. For the non-ELL students, letter identification explained 8% of the variance, rhyme detection accounted for an additional 5%, oral cloze represented 3%, and rapid naming explained 2% of the variance. For the ELL students, rhyme detection explained 11% of the variance in outcome score, letter identification accounted for 7%, and rapid naming and oral cloze were not significant predictors. The authors arrived at several important conclusions including that limited English proficiency does not impede reading development. In fact, for the most part, ELL students with risk status in Kindergarten caught up to or surpassed their non-ELL counterparts on reading measures by the end of second grade. The results of the study provide evidence that the development of early reading skills in ELL students is very similar to how those skills develop in non-ELL students, and phonological awareness instruction can be provided to ELL students in English rather than relying on building skills in the primary language before transferring instruction to English.

Conclusion

Given the relatively high profile that early literacy instruction has received in recent years, there are surprisingly few empirical demonstrations of effective interventions for ELLs who struggle with reading (Klingner, Artiles, & Mendez Barletta, 2006). The review above highlights a few studies with salient findings that contribute to
this knowledge base. Although there is a literature base that supports an alternative and qualitatively different approach to instruction for ELLs, it is not an empirical one (Gersten, Baker, Haagar, & Graves, 2005). This non-empirical base is in contrast to the studies outlined above and the emerging consensus that ELLs acquire reading skills in the same fashion as non-ELLs, therefore many of the same principles of effective instruction apply. Specifically, Gersten, Baker, Haagar, and Graves point out that the amount and quality of explicit instruction in phonics, phonemic awareness, and vocabulary were related to the level of reading proficiency demonstrated by ELLs and that students who received this type of instruction reached performance levels similar to non-ELLs. In an interview with Chamberlain (2006), Sharon Vaughn summarizes this issue this way, “I can say that when English language learners who are at risk for reading problems are provided intervention that matches the language of their core reading instruction (in our studies Spanish or English, depending on the group), when they are provided specific and explicit instruction that maps onto their instructional needs, that they make progress that is equivalent to, if not exceeding, the progress of monolingual English students” (p. 170).

Elements of Treatment Delivery

Once students are identified, and an effective intervention with the characteristics described earlier is selected, treatment delivery becomes a critical factor in ensuring its success. In order for supplemental instruction to be effective enough to improve student outcomes, the treatment needs to be of sufficient length in minutes, consistently implemented, and remain in place for a sufficient number of weeks over the school year (Lyon, 1993). After all, it is not only important that the student gets treatment, but also that the treatment is in place long enough to be effective. Similarly, once a treatment is in
place, a high level of implementation integrity should be maintained. This section addresses the literature on treatment integrity and treatment length or treatment duration.

Treatment Integrity

Treatment integrity refers to the extent to which an independent variable is implemented as intended (Cooper, Heron, & Heward, 1987). Treatment integrity is a critical component of intervention research because the conclusions drawn about the effects of an applied treatment are predicated upon the assumption that the intervention described is indeed the one that took place. Whereas reliability measures are used to ensure accuracy of measurement of the dependent variables, treatment integrity measures the accuracy and consistency of the independent variables. Ensuring a high level of treatment integrity is vital to the validity of the experiment. Internal validity refers to the extent to which changes in the dependent variables are attributable to the application of the independent variable and, “without documentation as to how well the intervention was implemented, it is impossible to attribute the observed changes to the independent variable” (Schlosser, 2002, p. 37). Furthermore, in single subject research, the primary method of strengthening the external validity of some treatment is through building a replicative history by conducting systematic and direct replication of some experiment. However, without a clear operational definition, which is evaluated with systematic and repeated integrity checks, building a replicative history is virtually impossible. One cannot replicate what was never defined or evaluated. It is particularly important that treatment integrity is assessed and reported in single subject research since this method relies on the demonstration of a functional relationship between the independent and
dependent variables and the meaning of that relationship is based upon the faithful and
consistent application of the independent variable, which lends credibility to the results.

Given the importance of ensuring the integrity of the independent variable, some
authors have noted an absence in the literature that has been called a “curious double
standard” (Peterson, Homer, & Wonderlich, 1982). Although operational definitions and
reliability data are provided almost invariably for the dependent variables, seldom are the
same standards applied to the independent variables. Peterson and colleagues (1982)
conducted a review of the experimental articles published in the *Journal of Applied
Behavior Analysis* from its inception to 1980. The authors found 36% of the articles
published data referring to the assessment of the independent variable. When Gresham,
Gansle, and Noell (1993) reviewed the same literature but limited the articles to
educational interventions for children, they found that 15.8% of the published articles
reported the integrity of the independent variable. Furthermore, in this 12-year span, no
trend toward inclusion of these valuable data was observed. In a more recent review of
the literature in learning disabilities, Gresham, MacMillan, Beebe-Frankenberger, and
Bocian (2000) evaluated three major journals in learning disabilities research: *Journal of
Learning Disabilities, Learning Disability Quarterly, and Learning Disabilities Research
& Practice*. Between January 1995 and August 1999, 469 articles were published in these
journals and 13.6% (65) were intervention research. Of these articles, only 12 (81.5%)
measured and reported data on the integrity of the independent variable. The authors
concluded that little can be discerned from the studies published in that time period and
the special education community is left with two outstanding questions: (a) How are
intervention treatments being implemented, and (b) what is the nature of the relationship
between treatment integrity and treatment outcomes? The seriousness of this gap in the intervention literature cannot be overstated and some have even suggested that the inclusion of treatment integrity data be given considerable weight in the publication decision (Schlosser, 2002).

Recently a series of articles in *Clinical Psychology: Science and Practice* addressed the question of the relationship between treatment integrity and treatment outcomes. Perepletchikova and Kazdin (2005) noted that although some literature tends to suggest a positive relationship between treatment integrity levels and treatment outcomes (Gresham, 1989; Peterson, Homer, & Wonderlich, 1982), a breadth of literature exists that does not support this hypothesis (Bein et al., 2000; Burke, 1996; Noell, Witt, Gilbertson, Ranier, & Freeland, 1997; Patton, 1988; Toffolo, 2000). In fact, the assumption that high levels of treatment integrity produce improved outcomes is simply not founded; indeed this relationship appears to be rather inconsistent and complex. Perepletchikova and Kazdin point out that three kinds of treatment integrity actually exist: treatment adherence, therapist competence, and treatment differentiation. Treatment adherence refers to level at which the interventionist adhered to specified procedures in the treatment protocol. Competence implies the sensitivity and timeliness of how strategies are applied to a particular problem. Differentiation is the extent to which treatments under investigation were sufficiently different to be salient.

The authors propose several factors that can contribute to the inconsistency of the relationship between treatment integrity and outcomes. Among these factors are erroneous self-report of integrity levels, rather than direct observation by a third party, a significant lack of empirical demonstrations in which integrity levels are intentionally
manipulated, and the possibility that in some cases, deviations from the treatment protocol can actually augment procedures with more efficient or timely techniques. In this instance, low integrity levels are not indicative of less effective treatment techniques. Furthermore, Perepletchikova and Kazdin contend that treatment integrity may serve only as a proxy variable for other, more important, variables that affect treatment outcomes. Interventionist, treatment, and client characteristics were all mentioned as possible factors related to treatment outcomes. For example, interventionists with substantial experience may be more difficult to adapt to new treatment protocols and could contaminate the integrity by incorporating other techniques. As previously mentioned, this can affect treatment outcomes positively or negatively, but will definitely negatively impact the level of treatment integrity. Client, or student, characteristics can also affect integrity levels. If a student is particularly difficult to manage or presents intense problems, the interventionist may become discouraged or may lack motivation to address such a serious case. Rather than putting in the effort to remediate the student’s difficulties, the interventionist may instead refer the student for other services (special education) because he or she perceives the case as too difficult to manage. Especially when students present serious deficits, interventionists become less motivated to work with the students because very little success is yielded for considerable time and effort. Finally, characteristics of the treatment itself could affect the integrity with which it is implemented. Treatments that are too complex or time consuming are less likely to be implemented exactly as intended, especially those that require prior planning or extensive preparation.
In order to increase interventionist competence, and thereby support the likelihood of improved integrity, Perepletchikova and Kazdin (2005) recommend direct training with modeling, role-playing, feedback, rehearsal, and booster sessions, if needed. The authors note that continued supervision and feedback are just as important as initial training. Additional recommendations include: (a) treatment manuals that include the rationale for instruction, scripted instructions, instructions for sequencing, and procedures for addressing deviations, (b) explicit supervision methods that enhance accuracy of implementation, and (c) informing interventionists that the purpose of observations is to assess integrity. Finally, although most researchers assess and report overall integrity, which is usually a composite score of each session’s integrity, component and session scores should also be provided. This may provide more valuable information about the aspects of integrity affecting outcomes. The session scores can be reported as a range scores, whereas component scores report the level of implementation for each intervention component described on the assessment form. Specifying component scores may permit more fine-grained analysis into which components exert more influence over outcome scores, which will advance the knowledge base on the relationship between treatment integrity and outcomes.

*Duration of Treatment*

Although recent legislation such as the No Child Left Behind act mandates low-performing schools to provide supplementary instructional services for struggling readers, financial and personnel resources remain limited. Furthermore, mandates for supplementary services do not have explicit requirements for treatment duration. The supplemental intervention literature represents broad variability in the application of
these mandates. The intensity of instruction can range from individual to small and larger groups and between 1 and 5 days per week. Furthermore, session duration may be anywhere between 10 and 50 minutes. Given the growing literature base on supplemental early reading intervention, it is surprising that very few studies have systematically investigated the effects of treatment duration on outcomes (Al Otaiba, Schatschneider, & Silverman, 2005). The relevance of this question may become more important to educators and administrators as the need for more services, with fewer resources, is apparent at the school level. If more is known about the optimal treatment durations for specific reading difficulties, resources can be used more efficiently. The following is a review of two recent studies that investigated the relationship between treatment duration and outcomes.

In a three-year study of 49 middle school students with learning disabilities, Mercer, Campbell, Miller, Mercer, & Lane (2000) used a three-group pretest posttest group design to evaluate the effects of different levels of a fluency-building intervention. The levels of the intervention were based on three categories for duration of treatment: 6-9 months, 10-18 months or, 19-25 months. Students were given a phonics page, a sight words page, and a story from the Great Leaps (Campbell, 1995) curriculum to read repeatedly with a teacher or instructional assistant. The authors hypothesized that the additional phonics and sight word instruction would enhance the effects of repeated readings. The passages were timed for one-minute and reading fluency was scored for pauses of greater than three seconds, substitutions, omissions, and mispronunciations. All instructional sessions were one-on-one for seven minutes daily. Since the middle school’s
basal reading program was considered too difficult for these students, the fluency intervention was the only reading instruction provided during the intervention period.

The results suggest that all three groups made substantial fluency gains. The effect size for the fluency curriculum based assessments were 2.42, 1.52, and 1.55 for the 19-15 months, 10-18 months, and 6-9 months groups respectively. Although all groups made progress, it appears that the rate of progress was steeper for the 6-9 month group. These students gained almost two years on the *Great Leaps* reading assessment. The 10-18 months and 19-25 months groups gained just over three years on these assessments over a longer period of time. The authors compared the growth rates for the 6-9 months and 19-25 months groups and concluded that the differences may be due to several factors. First, students assigned to the 19-25 months group had the most severe disabilities. All students in this group began at the pre-primer level, whereas less than 50% of the 6-9 months group began at this level. It is possible that since this group had the weakest readers, they were slower to make gains. Another possibility is that a novelty effect may have influenced performance. The authors note that performance in an intervention is often strongest at the beginning and thus, the longer one is in intervention, the more gradual the progress. Finally, it should be taken into consideration that this study was conducted with middle school students. The greatest reading growth occurs between first and second grades. Since, the 6-9 months group pretested in this range (1.5), the match between intervention strategy and reading readiness may have been more optimal than for others. The findings of this study suggest that for certain students, those who need some support but perhaps do not require the most intensive intervention, a shorter treatment duration may be sufficient to ameliorate their deficits. Overall, this
study is consistent with the literature on building oral reading fluency, and expands the knowledge base by demonstrating a differential rate of growth among three different treatment durations.

Al Otaiba, Schatschneider, and Silverman (2005) examined the effects of a tutoring intervention for early reading skills. Twelve Kindergarten teachers from four high poverty urban schools agreed to participate in this study. Students were pretested with the DIBELS Letter Naming Fluency (LNF) and Initial Sound Fluency (ISF) subtests. Students who scored fewer than three correct initial sounds and named fewer than two letters were considered at risk and eligible for this intervention. A total of 73 students participated in the Tutor Assisted Intensive Learning Strategies (TAILS) program. TAILS is based on the Peer-Assisted Learning Strategies for Kindergarten (K-PALS; Fuchs et al., 2001a), which has been shown to improve phonological awareness and spelling with Kindergarten students with and without disabilities in Title I and non-Title I schools. Activities in the TAILS program include phonemic awareness exercises, phonics, fluency, vocabulary, and comprehension. An example of the phonemic awareness activity includes using manipulatives to represent sounds in speech and teaching students increasing smaller units, from syllables to onsets and rimes, and finally phonemes. Other exercises include word building and decoding using a model-lead-test format.

Students were randomly assigned to three levels of the intervention: (a) TAILS 4 days per week, (b) TAILS 2 days per week, or (c) control in order to investigate the effects of differing instructional intensities, since there is a lack of literature that explicitly manipulates this aspect (Wasik, 1998). TAILS sessions lasted 30 minutes and
varied in the number of days by condition. Students in the control condition were read to by a tutor twice per week for approximately 20 minutes. The authors concluded that students in the 4-day condition made statistically significant greater gains than students in either the 2-day or control groups. Students in the 4-day groups demonstrated effects sizes of .79, .90, and .83 for Word Identification, Passage Comprehension, and basic skills on the Woodcock Johnson Reading Mastery test (Woodcock, 1987). Students in the 2-day group generally performed better on outcome measures than controls, however, the difference was not statistically significant. This study contributes to the literature by demonstrating how different levels of treatment duration, or overall instructional intensity, can affect reading outcomes for Kindergarten students at risk for reading failure in a high poverty urban school. In this case, more consistent and frequent intervention produced better outcomes.

Summary

Many urban elementary schools find that students who enter Kindergarten do not have the phonological skills to benefit from the general classroom instruction. These students require supplemental instruction to prepare them for subsequent reading instruction in Kindergarten and the early elementary grades. It is important that researchers take resource allocation into consideration and begin to systematically explore the possibility of not only establishing which interventions are effective, but also exactly how much is sufficient for students to benefit and maintain their skills. Additionally, Kjeldsen, Niemi, and Olofsson (2003) concluded that, “Over and above that sufficient dose, a careful training procedure seems to be more important for long-term training effects than the quantity of training. A similar finding regarding the importance
of the quality of the intervention was reported by Schneider et al. (1997, p. 335) who concluded that ‘a careful training procedure is necessary to secure substantial long-term effects of the program”’ (p. 362). Examination of the relationship between treatment duration and rate of progress with proven effective supplemental early reading interventions is a worthwhile direction for future research.

Conclusion

The importance of effectively teaching young children to read cannot be overstated. It is commonly noted that reading is the key that unlocks all other learning in formal schooling. Currently, most students will benefit from regular classroom instruction. However, for the remaining 20% of students, additional efforts may be required. It is estimated that in urban settings this figure may be even greater and effective supplemental intervention must be provided to a larger proportion of students. To exacerbate the situation, urban classrooms are faced with an increasingly diverse student body, including those who come from culturally and linguistically diverse backgrounds. Often, these students represent a wide variety of languages and it may be cost prohibitive for urban districts to provide enough multi-lingual instructors to provide intervention.

It is a worthwhile endeavor to approach this problem of educating diverse students at risk with fewer resources and personnel by examining the extent to which known principles of effective instruction for native English-speakers can be applied to ELLs and how effective the intervention can be when implemented not by the classroom teacher, but an instructional assistant. As discussed in this review, the guidelines for effective intervention are: (a) making the instruction more explicit, (b) increasing the
intensity of the instruction, and (c) providing sufficient support to maximize student performance. Several studies reviewed in this section have illustrated that when these principles are applied students, both ELL and non-ELL, can improve their reading skills. Furthermore, in some cases it has been demonstrated that paraprofessionals, or instructional assistants, can deliver these interventions effectively.

High quality training and support is needed for any professional or paraprofessional to implement interventions with integrity. And there is an intuitive connection between high levels of treatment integrity and improved student outcomes. However, the findings on this topic are still mixed and more research is needed to determine the effect that integrity exerts on outcomes. Given the scarcity of resources for training and supervision, it’s crucial that the literature base is clear on the importance of treatment integrity for student outcomes.

Related to this issue is the question regarding the optimal treatment duration. If sufficient treatment durations is not provided, students will not adequately benefit from the efforts. However, if there is no additive benefit of providing intervention beyond some period of time at which students will reliably maintain gains, these resources may be better spent by providing interventions for other students.
The report of the National Reading Panel has set some preliminary parameters by suggesting that results of their meta-analysis found that between 5 and 18 hours of instruction is the optimal treatment duration. That is, outcomes from studies that lasted less than five hours or more than 18 resulted in less robust outcomes compared with studies that provided durations between 5 and 18 hours. As it stands, many studies report the period of time over which intervention occurred, but do not specify the intervention’s duration in actual minutes or hours. This information may be useful to examine the covariance between treatment duration and student outcomes in more detail.
CHAPTER 3

METHOD

The methods used in this study are described in this chapter. Specifically, this section will describe the participants, settings, primary researcher, definitions and measurement of the dependent variables, definitions and measurement of the independent variables, general procedures including training, placement testing and intervention procedures, materials, experimental design, data analysis, and measurement of social validity.

Participants

The participants in this study were 61 treatment students, 32 control students, six Instructional Assistants (IAs), and two graduate students in special education. Student participants and IAs were from one of three elementary schools in the Columbus Public School system. Principals consented to the IAs’ participation in the study and signed consent forms were obtained for treatment students. The following describes the treatment students, control students, IAs, graduate student instructors, and researcher.
Screening Procedures

Treatment students were selected based on the results of screening for risk markers using the *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS; Good & Kaminski, 2002.), Woodcock Johnson Tests of Achievement III (WJ III; Woodcock, McGrew, & Mathur, 2001) Letter-Word Identification (LWID) and Word Attack (WA) subtests, and parental consent to participate.

**DIBELS.** Benchmark assessments are given in autumn, winter, and spring and result in one of three intervention recommendations: (a) benchmark (no risk), (b) strategic (limited intervention necessary), and (c) intensive (substantial intervention necessary). DIBELS benchmark assessments consist of between two and four subtests, which change throughout the year according to developmental reading targets. The autumn benchmark assessment covers (a) Letter Naming Fluency (LNF) and (b) Initial Sound Fluency (ISF). The spring benchmark assessments consist of (a) LNF, (b) Nonsense Word Fluency (NWF), and (c) Phoneme Segmentation Fluency (PSF). All Kindergarten students from the three participating schools were screened with the DIBELS and scores were entered into the DIBELS database, which provided intervention recommendations. All intensive and most strategic students were eligible for inclusion in the study. Due to limited group sizes and intervention times, not all strategic students were included. The decision was based on specific DIBELS scores and teacher recommendations.

**Woodcock Johnson Tests of Achievement III.** All Kindergarten students enrolled in each school were also given the LWID and WA subtests from the WJ III. The LWID subtest requires the student to identify or read a list of letters or words of increasing
difficulty. The student must read a list of increasingly difficult nonsense words on the WA subtest. A ceiling is reached and the test is concluded once the student has incorrectly responded to six consecutive items. Students who scored below grade level and had a strategic or intensive DIBELS recommendation were eligible for inclusion in the study.

_Parental consent._ A letter describing the study and an accompanying consent form (Appendices A and B) was sent to the parents or guardians of the students who met the above criteria. Consent forms for ELL students were translated into their native language (e.g. Somalian or Spanish; Appendices C and D, or E and F). Although all intensive students were provided intervention (as required by designation as a Reading First school under the No Child Left Behind act), only the data for students who returned the signed consent form were included in the study.

_Treatment Students_

Treatment students consisted of both English Language Learners (ELLs) and native English speakers (Non-ELLs). Intervention groups were homogeneous. That is, groups were either comprised of all ELLs or Non-ELLs. No group consisted of a mix of ELLs and Non-ELLs. Therefore, the groups will be treated separately.

_Non-ELL students._ Students who met all of the above inclusion criteria were entered in the study, resulting in 38 Non-ELL treatment participants. Of this group, 34.2% (13) were female and 65.7% (25) were male. African-American students accounted for 47.3% (18), 39.4% (15) were Caucasian, 7.8% (3) were Latino, and 5.2% (2) were multi-racial. The mean pretest scores for LWID and WA were 9.6 and 2.1 raw score points, respectively. On the fall DIBELS benchmark assessment, 65.7% (25) were
in need of strategic intervention, while 34.2% (13) were in need of intensive intervention. All of the students except two were free or reduced lunch status, an indicator of low income.

ELL students. A total of 23 ELLs participated in the study. Of this group, 39.1% (9) were female, 60.8% (14) were male, 82.6% (19) were Somalian, 13% (3) were Latino, and 1 was Asian. The mean pretest scores for LWID and WA were 6.4 and 1.5 raw score points, respectively. On the fall DIBELS benchmark assessment, 43.4% (10) were intensive, 47.8% (11) were in need of strategic intervention, and 8.6% (2) were benchmark, but included in the study to establish an additional group who placed at the same lesson in the curriculum. Three ELL students (13%) were in first grade, but the rest were in Kindergarten. All students in this group qualified for free or reduced lunch.

Control Students

Data were retained for 32 students who did not receive intervention in order to address the first research question pertaining to the amount of variance in outcome scores that could be attributed to participation in the intervention. It is important to note that students were neither randomly selected nor assigned and the control students are not comparable to the treatment students since students were only selected to participate in the study due to deficits in early reading indicators. Seven control students were selected from School 1, 17 attended School 2, and eight attended School 3. Of this group, males accounted for 62.5% (20), and 37.5% (12) were female, 56.3% (18) were African American, 31.3% (10) were Caucasian, 6.3% (2) were Latino, 3.1% (1) was Asian, and 3.1% (1) was multi-racial. The mean pretest scores for LWID and WA were 14.3 and 9.7, respectively. The fall DIBELS benchmark was administered to each control student. Four
students (12.5%) needed strategic intervention, but did not test below grade level on the WJ-III and therefore were not included in the treatment group. The remaining control students (28) were at benchmark on the fall assessment. All students in this group except one qualified for free or reduced lunch.

<table>
<thead>
<tr>
<th></th>
<th>Treatment non-ELL</th>
<th>Treatment ELL</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>65.7%</td>
<td>60.8%</td>
<td>62.5%</td>
</tr>
<tr>
<td>Female</td>
<td>34.2%</td>
<td>39.1%</td>
<td>37.5%</td>
</tr>
<tr>
<td>African American</td>
<td>47.3%</td>
<td>-</td>
<td>56.3%</td>
</tr>
<tr>
<td>Caucasian</td>
<td>39.4%</td>
<td>-</td>
<td>31.3%</td>
</tr>
<tr>
<td>Latino</td>
<td>7.8%</td>
<td>13.0%</td>
<td>6.3%</td>
</tr>
<tr>
<td>Multi-racial</td>
<td>5.2%</td>
<td>-</td>
<td>3.1%</td>
</tr>
<tr>
<td>Somali</td>
<td>-</td>
<td>82.6%</td>
<td>-</td>
</tr>
<tr>
<td>Asian</td>
<td>-</td>
<td>4.4%</td>
<td>3.1%</td>
</tr>
<tr>
<td>Free/reduced lunch</td>
<td>94.7%</td>
<td>100%</td>
<td>96.9%</td>
</tr>
<tr>
<td>DIBELS intensive</td>
<td>34.2%</td>
<td>43.4%</td>
<td>-</td>
</tr>
<tr>
<td>DIBELS strategic</td>
<td>65.7%</td>
<td>47.8%</td>
<td>12.5%</td>
</tr>
<tr>
<td>DIBELS benchmark</td>
<td>-</td>
<td>8.6%</td>
<td>87.5%</td>
</tr>
<tr>
<td>LWID pretest</td>
<td>9.6</td>
<td>6.4</td>
<td>14.3</td>
</tr>
<tr>
<td>WA pretest</td>
<td>2.1</td>
<td>1.5</td>
<td>9.7</td>
</tr>
</tbody>
</table>

Table 3.1: Demographic information and pretest scores for treatment and control students.
**Instructional Assistants**

Six instructional assistants (IAs) delivered the intervention to the non-ELL treatment students. Consent to participate in the study was secured through written consent of the school principal (see Appendix G).

**School 1.** Two IAs from School 1 participated. D.C. was an African-American male with eight years of experience as an IA, however, this project represented the first time he taught a structured curriculum in a small group setting. The amount of formal education D.C. had is not known. D.C.’s instructional group consisted of five students. D.M. was an African-American female with 10 years experience as an IA. She participated in a similar project during the previous school year and co-taught the curriculum with a doctoral candidate who was completing her dissertation. Additionally, D.M. has completed several university level teacher education courses, but has not yet completed an education degree. D.M. taught five students in her small group.

**School 2.** Two IAs and one GA from School 2 participated. A.H. was a Caucasian male who was employed by the district as a substitute IA for two years prior to the onset of the study. The 2005-2006 school year was the first year A.H. was employed as a fulltime IA. This was also his first experience teaching reading in a small group format. He held a bachelors degree in a non-education related field. L.P. was also an IA at School 2. She was a Caucasian female with 20 years of experience as an IA or classroom volunteer. She had informal experience tutoring students one on one or working with her own children at home. L.P. did not have any university-level credits. Both L.P. and A.H. taught three small groups five days per week. The small groups consisted of between two and four students.
One graduate assistant (GA) in special education provided instruction at School 2. L.K. was pursuing her masters of arts in education at The Ohio State University and this project served as her thesis study. She worked with three small groups of ELLs in a multiple baseline design and taught a total of 11 students. L.K. received her bachelor’s degree in early childhood education from Miami University in 2001. She has experience as both a permanent substitute in an elementary school and a preschool and second grade teacher for the Department of Defense in Okinawa, Japan. In that position she taught a diverse group of children, some of whom were ELLs who primarily spoke Spanish or Japanese.

School 3. Two IAs and one GA provided instruction for participants in School 3. R.O. was a Caucasian female who had been employed as an IA with the district for five years. She did not have any previous experience with delivering instruction to students. Her job was primarily to fill in as needed throughout the school. For example, R.O. frequently assisted with book fairs, secretarial duties, or chaperoning field trips. S.W. was also an IA at School 2. She was an African-American female with five years of experience in this district. S.W. reported that she had experience as a safety net reading teacher. In this position she delivered reading interventions in small groups or individually for students who experienced reading difficulties. The amount of formal education these IAs had is not known.

One GA in special education delivered instruction for three groups of ELL students. M.C. was pursuing a master’s in education at The Ohio State University. She held bachelor’s degree in psychology and had no prior teaching experience. M.C. delivered instruction for 12 students who were divided into three small groups.
Table 3.2: Background information for the instructors

<table>
<thead>
<tr>
<th>Instructor</th>
<th>School</th>
<th>Position</th>
<th>Years Exp.</th>
<th>Gender</th>
<th>Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>D.C.</td>
<td>1</td>
<td>IA</td>
<td>8</td>
<td>M</td>
<td>AA</td>
</tr>
<tr>
<td>D.M.</td>
<td>1</td>
<td>IA</td>
<td>10</td>
<td>F</td>
<td>AA</td>
</tr>
<tr>
<td>A.H.</td>
<td>2</td>
<td>IA</td>
<td>2</td>
<td>M</td>
<td>C</td>
</tr>
<tr>
<td>L.P.</td>
<td>2</td>
<td>IA</td>
<td>20</td>
<td>F</td>
<td>C</td>
</tr>
<tr>
<td>L.K.</td>
<td>2</td>
<td>GA</td>
<td>5</td>
<td>F</td>
<td>C</td>
</tr>
<tr>
<td>R.O.</td>
<td>3</td>
<td>IA</td>
<td>5</td>
<td>F</td>
<td>C</td>
</tr>
<tr>
<td>S.W.</td>
<td>3</td>
<td>IA</td>
<td>5</td>
<td>F</td>
<td>AA</td>
</tr>
<tr>
<td>M.C.</td>
<td>3</td>
<td>GA</td>
<td>0</td>
<td>F</td>
<td>C</td>
</tr>
</tbody>
</table>

Settings

This study was conducted in three elementary schools in an urban school district consisting of 63,098 students, 4,339 of which were ELLs, in 153 schools. The district employed 3,838.4 teachers and reported a student/teacher ratio of 16.4:1. This project was supported in part by the U.S. Department of Education, which funded the model demonstration grant “Culturally Responsive Instruction for African American Males: Reducing Disciplinary and SED Referrals”. Additionally, this project was partially supported by a grant from Ohio State University’s P-12 program. The former project has taken place exclusively in School 1 for the past four years. The 2005-2006 school year was the first year this project was expanded to Schools 2 and 3.

School 1

This school’s enrollment was 238 students in grades preschool through fifth. African-American students accounted for 69% (164) of the total enrollment, Caucasians
accounted for 24.4% (52), 3.8% (8) of students were Latino, 2% (4) were Asian, and 0.8% (2) were Multiracial. No students of Somalian decent were enrolled in the Kindergarten at this school. Students eligible for free or reduced lunch accounted for 69% (166) of the total enrollment. Special education services were provided for 22% (52) of the total enrollment.

School 1 used the LACES (Literacy Across Columbus Elementary Schools) reading program. LACES is a district-designed reading program based on data from students in the district, the recommendations of the National Reading Panel, and the Ohio Academic Content Standards for reading and language arts. The reading program includes activities on phonemic awareness, phonics, word study, fluency, writing, spelling, and comprehension. LACES was implemented every morning from 9:30 to 11:30.

HOSTS, a university-community partnership tutoring program, took place every Tuesday and Thursday for 30 minutes. Students were paired with a tutor from the community who read to the students with the purpose of instilling an appreciation of reading and helping to address reading deficiencies. A total of 200 adults who were given release time from their employer volunteered their time for this program.

School 2

The total enrollment for School 2 was 336 students in grades preschool through fifth. Sixty-two percent (207) were African-American, which included Somalians. Thirty-three percent (112) were Caucasian, 2% (7) were Latino, 1% (2) were Asian, and 2% (7) were Multiracial.
Students eligible for free or reduced lunch accounted for 96.4% (324) of the total student enrollment. Special education services were provided for 15.4% (52) of the students.

School 2 used the *Trophies* reading program (Beck, Farr, & Strickland, 2003) from 9:30 to 11:30 every morning. Trophies is a comprehensive reading program that has activities on direct and guided reading instruction, phonics, phonemic awareness, writing, vocabulary, fluency, and comprehension. Trophies is used in 2,000 school districts across the United States. It includes a supplemental intervention program, an ELL resource kit, and a parallel Spanish-language version of the program, *Trofeos*. The *HOSTS* program took place every Tuesday and Thursday.

*School 3*

The total enrollment for School 3 was 234 students in grades preschool through fifth. Fifty-four percent (126) were African-American, which included Somalians. Twenty-eight percent (66) were Caucasian, 12.3% (29) were Latino, 2.9% (7) were Asian, and 1.7% (4) were Multiracial. Students eligible for free or reduced lunch accounted for 94% (220) of the total enrollment. Special education services were provided for 13.6% (32) of the students. School 3 also used the *Trophies* program from 9:30 to 11:30 every morning. The *HOSTS* program was not implemented in this school.

*Instructional settings*

*School 1.* Both IAs, D.C. and D.M., conducted the intervention from approximately 9:15 to 9:45 on Mondays, Wednesdays, and Fridays. D.C. provided instruction at a 5’ X 2’ rectangle table in the teacher’s lounge. D.C. sat on one side of the table and the five students sat across from him and at the sides of the table. The teacher’s
lounge was a small, quiet room with kitchen equipment that was not in use during the intervention period. D.M. conducted the intervention for five students at a round table (5’ diameter) in the hallway outside the preschool classroom. This was the quietest area available at this time.

School 2. The interventions were conducted between 10:45 and 12:15. Each IA conducted three groups for 30 minutes each five days per week. A.H. conducted his groups at a 5’ X 2’ rectangular table in a vacant classroom on Mondays, Wednesdays, and Fridays. A.H. sat on one side of the table and the students sat across from him. On Tuesdays and Thursdays, HOSTS took place in this classroom, so A.H. conducted his groups in the back corner of a Kindergarten teacher’s classroom on these days. L.P. conducted the intervention in the corner of the same teacher’s classroom each day. She sat on one side of the 5’ X 2’ rectangular table and students sat across from her. L.K. took her students to a small supply room that had a 5’ diameter round table. She and the students sat around the table during the intervention.

School 3. Both IAs conducted intervention groups from 11:00 to 11:30 each day. R.O. worked with students in a quiet, partitioned area of one of the Kindergarten classrooms. R.O. sat at a 5’ X 2’ rectangular table and the students sat across from her. R.O. conducted another group from 2:15 to 2:45 each day in the same area. S.W. conducted her intervention group in a quiet corner of the school library at a 5’ diameter round table. M.C. also worked with students in the same area of the library, which was partitioned using tall bookshelves. However, for the last four weeks of the study, M.C. moved her groups into the other Kindergarten teacher’s classroom. M.C. conducted the intervention at a 5’ X 2’ rectangular table with the students sitting across from her.
Primary Researcher and Secondary Observer

The primary experimenter was a doctoral candidate in special education and applied behavior analysis at The Ohio State University. She completed her bachelor’s degree in music education at Kent State University in Kent, Ohio in 1999. Upon completing her bachelor’s degree she was hired as a teacher in a private juvenile correctional and rehabilitation facility in northeast Ohio. This facility served adjudicated youth who were placed by the juvenile court system for a minimum commitment. However, once the duration of the commitment was fulfilled, facility staff, psychologists, probation officers, social workers, teachers, and other professionals reviewed each case to determine if the youth could be returned to the community, depending on the offense committed. Although the facility serviced both male and female offenders, the primary researcher taught only males. She taught males in sixth through eighth grades for the first six months and then taught males in ninth through twelfth grades for the remainder of her two years of employment. Responsibilities included teaching all core subject areas as well as two elective subjects such as health, physical education, or music theory each period. Additionally, a primary responsibility of this position was establishing classroom management and structure as well as teaching and monitoring appropriate social behaviors. The primary researcher was trained in physical restraint and obtained a non-tax certificate through the state of Ohio for the duration of her employment. She also attended related professional development in-services. In 2001 she entered graduate school at The Ohio State University and completed her master’s of the arts in special education (mild to moderate disabilities) in 2003. That same year she entered the doctoral program in special education and applied behavior analysis. She has been conducting or
assisting with academic and behavioral intervention studies, teacher in-services, and seminars at School 1, which was the site of a model demonstration grant, since 2002.

A secondary observer participated in this study. She was a doctoral student in special education and applied behavior analysis at The Ohio State University and also completed her master’s degree at the same university. She served as a secondary observer for probes of implementer integrity.

Definition and Measurement of the Dependent Variables

The primary dependent variables were two subtests from the Woodcock Johnson Tests of Achievement – III (WJ III) and the DIBELS benchmark assessments. The DIBELS benchmark assessments were given three times during the year: fall, winter, and spring, but for the purposes of this study the results from only the fall and spring assessments were used. Due to designation as a Reading First school, Schools 2 and 3 were required to administer the DIBELS benchmark assessments. These data were obtained from the schools’ data managers. The primary researcher administered the benchmark assessments for School 1. The WJ III subtests were given in fall and spring by the primary researcher and two graduate assistants in special education. The assessments are described in more detail next.

*Woodcock Johnson III Letter-Word Identification and Word Attack Subtests*

All Kindergarten students were administered both subtests in the fall for risk verification and selection for participation in the study. In the spring the subtests were given to all treatment and control students as a posttest measure. The Letter-Word Identification (LWID) subtest contains 79 items and measures letter and word recognition. The student is shown a page of letters or words and asked to point to or say
certain letters or words. For ages 5 to 19 this subtest has a median reliability of .91. The Word Attack (WA) subtest contains 30 items and measures skill in applying letter sound knowledge to nonsense words. The first three items require the subject to identify singular letter sounds. The remainder of the items are nonsense words such as /tiff/ and /zoop/. For ages 5 to 19 this subtest has a median reliability of .87. Students were taken to a quiet area, usually in the hall outside the classroom, and administered the test individually. The students were told that they were going to be asked some questions. Some may be easy and others may be hard, but to try their best. A ceiling was reached when the student answered six consecutive items incorrectly and the test was stopped. The average time to deliver this assessment was approximately three minutes per student in the fall and five to 10 minutes per student in the spring.

**DIBELS Benchmark Assessments**

The DIBELS assessments are standardized one-minute fluency measures that result in a recommendation for the level of intervention services the student needs. The DIBELS should be administered three times per year in order to monitor student progress and response to intervention. In the fall the Letter Naming Fluency (LNF) and Initial Sound Fluency (ISF) subtests are administered to Kindergarten students. ISF reflects the student’s level of phonological awareness by recognizing and producing the initial sounds in words. Alternate form reliability on this measure is .72 and when the measure is repeated four times, the resulting mean is .91. On the LNF subtest students read as many upper and lowercase letters as they can in one minute. Alternate form reliability is .88.
On the winter DIBELS benchmark assessment, Phoneme Segmentation Fluency (PSF) and Nonsense Word Fluency (NWF) are added. PSF requires students to individually produce the phonemes in an orally presented word. For example, the administrator of the test says “sat” and the student should respond with three clear, individual phonemes, “/s/ /a/ /t/”. Two week alternate form reliability is .88 and one-month reliability is .79. On the NWF subtest students are presented with a page of nonsense words that include consonant-vowel-consonant words such as “vif” and vowel-consonant words such as “ov”. Students are required to produce each individual sound or read the nonsense word in its entirety. This subtest measures the student’s level of mastery of letter-sound correspondence. This subtest has an alternate form reliability of .83. The spring benchmark assessment includes LNF, PSF, and NWF, but does not include ISF.

Definitions and Measurement of the Independent Variables

The following section will describe the Early Reading Intervention (ERI; Simmons & Kame’eu, 2003), and the methods used to measure the treatment quality and duration. These three variables were regressed on the dependent variables in this study.

Early Reading Intervention

The ERI is a supplemental reading intervention program with a high degree of explicitness on code emphasis skills that are necessary for developing phonological awareness. This curriculum is a byproduct of a five-year longitudinal study, Project Optimize, which was described in chapter 2. Simmons and Kame’enui were both part of the assessment group for the Reading First Initiative for the U.S. Department of Education. The results of Project Optimize (Simmons et al., 2003) suggest that 97% of
the students who received ERI instruction demonstrated faster achievement in phonological awareness and alphabetic principle skills than students who received other interventions. The program contains 126 lessons and each lesson can be taught in either one 30 minute period or two 15 minute periods. The first 15 minutes is dedicated to phonological awareness skills, alphabetic understanding, and word reading. Further phonological awareness activities are conducted in the remaining 15 minutes, along with writing and orthographical skills (connecting letter-sound knowledge and whole word writing). Additionally, this program comes with a comprehensive assessment to measure response to intervention. A placement test with directions for administration, four progress monitoring tests and checklists, and an exit test can be found in the assessment booklet. Before instruction begins, each student should be given the placement test, which evaluates skill in letter and sound naming, initial sound isolation, segmentation, letter sound correspondence, and regular word reading. The results of this assessment indicate where in the program is the most appropriate starting place for the student’s skill level. Once instruction has begun and the end of one of the four sections is reached, students should be given the progress monitoring test to verify that the student is prepared to move on to the next section. Additionally, a student progress checklist should be completed for each student every six lessons. This informal checklist keeps record of the skills each student has mastered and will alert the instructor to any difficulties students are having that may require more intensive instruction. Once the student reaches the last lesson in the program, an exit test is given to determine whether or not the student has mastered skills such as letter names and sounds, initial sound isolation, segmentation, letter-sound correspondences, and regular and irregular word reading.
The ERI comes with all the materials needed to successfully conduct the intervention. Four teacher’s guides contain the scripts for each of the 126 lessons. Teacher resource packages, the assessment book, Diz the Dinosaur puppet, Diz storybooks, write-on wipe off cards, and student activity books are contained within the kit. The manipulatives for each of the lessons consist of alphabet cards, smaller letter and word cards, letter tiles, picture cards, game boards, and two- and three-square strips. Two videos, a message from the authors and a demonstration video, come with each kit. All materials are contained within a colorful, sturdy cardboard box on which Diz the Dinosaur sits atop.

The first page of each lesson serves as an advance organizer for the activities to follow. The authors recommend that approximately 15 minutes will be necessary to prepare for each lesson, though that preparation time should decrease once the instructor becomes more familiar with the format. The lesson overview has three columns with approximately seven rows, which describe each activity. This first column describes the activity (e.g. alphabetic, writer’s warm-up, phonologic, integrated phonologic/spelling etc.). The middle column describes the materials needed for that activity. The last column indicates the amount of time that should be spent on each activity. Although the content and difficulty level vary among the four program parts, a general format prevails throughout the program. Furthermore, a model-lead-test paradigm is used to deliver the instruction. Every skill that students learn is first explicitly modeled by the instructor. If necessary, it can be modeled several times until students can attempt the skill while being led by the instructor. Finally, students are tested by performing the target skill alone. If errors occur, the sequence is repeated.
A typical ERI lesson from Part 1 began by introducing a letter and its sound. For example, the teacher read from the script, “The name of this letter is *m*. Say the name of the letter with me: *m*. What is the name of this letter?” Next the instructor said, “The sound for this letter is /mmm/. When you say /mmm/ your lips come together. Put your lips together and say /mmm/. Say the sound with me: /mmm/. What is the sound for this letter?” This introductory activity took approximately two to three minutes.

The objective for the next activity was to isolate the initial sound. The instructor used several picture cards depicting images that began with the target letter and one or two that did not begin with the target letter as non-examples. First, each picture was introduced so there was no confusion about the name of the picture. The instructor said, “This is man. What is this? This is mouse. What is this? This is map. What is this? This is fin. What is this?” Students responded with the picture name each time. Next, the instructor will modeled the task by saying, “I’ll say the picture name and tell whether it begins with the first sound /mmm/. *Man* has the first sound /mmm/. My lips come together when I say /mmm/, *man*. Next picture is *fin*. *Fin* does not have the first sound /mmm/.” In this case, the teacher demonstrated an example and a non-example. Finally, the instructor provided turns for each student to do alone. This activity required approximately six to seven minutes.

The following activity reintroduced the target letter name and sound. The format was the same as in the first activity. Next was the writer’s warm-up activity. Writer warm-up consumable activity sheets were provided for each student. The instructor held up the tracing card with the target letter and told the students to watch as he or she finger traced the letter. Next the instructor modeled writing the letter with an erasable marker.
The students then traced the letters with fingers and then wrote with pencils. The instructor gave feedback and guidance as necessary. Students were finally asked to circle their best letter. This activity took between two and three minutes.

Next students connected the sound /mmm/ to the letter m. Pictures that began with the target letter were introduced and placed in front of the students along with the m letter card and two distracter cards. The instructor modeled the task by saying the picture name and selecting the letter that matches the picture’s first sound. Each student was given a turn to demonstrate the same skill. Each student was also asked to give the letter name and sound each time a response was made. This activity took between two and three minutes to complete.

Finally, the letter cross-out game concluded the first lesson on the letter m. Each student was given a sheet that depicts several letter m’s and distracters. The instructor modeled finding the letter that makes the /mmm/ sound and crossing it out with a pencil. Next, students were given the opportunity to find m’s on their own and cross them out. Each time a student found and crossed out an m, the instructor asked the students to say the name and sound for the letter. This activity took between six and eight minutes to complete. In the case that errors were made, an “ongoing assessment” box appeared on each page that guided the instructor on the exact procedure for correcting errors. Appendix H depicts a sample ERI lesson.

Treatment Quality

Treatment integrity, also known as procedural integrity or procedural reliability, is established when the researcher collects data to verify that a given treatment has been implemented as intended. Treatment integrity data provides evidence on the treatment
implementation that may lend credibility to the observed effects. Often, in educational intervention research, treatment integrity data are collected by directly observing the intervention and keeping record of whether or not important characteristics of the treatment were observed. A checklist of treatment components is created for this purpose. However, a checklist that only indicates whether or not the treatment component was present is not a very sensitive measure since the component may occur only once and it must be scored as an occurrence even though this component may not have occurred on subsequent opportunities. This can happen frequently in educational research, and specifically with the ERI intervention. The first component on the observation checklist is “follows the script.” Obviously, it is difficult to score this important component as either an occurrence or a non-occurrence. In reality, instructors follow the script to varying degrees and this adherence to or divergence from protocol may affect the effectiveness of the treatment.

In order to build upon the study by Musti (2005), a similar treatment integrity checklist was used to evaluate implementation, but a quality scale was added (Appendix I). Seven components were scored as either: never occurred (0), rarely occurred (1), sometimes occurred (2), or always occurred (3). Each IA or GA was observed during the ERI instruction for 10 minutes once to twice weekly. During the observation, the primary researcher sat in an unobtrusive area that was within earshot and all materials and student behavior could be clearly seen. If the IA or GA asked for assistance during instruction, the primary researcher responded to the request. The primary researcher did not initiate feedback or offer assistance during instruction. However, when the lesson was finished, the primary researcher offered guidance and feedback as time permitted. Due to
scheduling conflicts, a predetermined meeting for performance feedback was not established. At the conclusion of the observation, the primary researcher tallied the quality scores. A possible 21 points could be earned for each observation. IAs and GAs were not told what their scores were. A mean treatment quality score was determined for each group that instructor led. For example, if an instructor had more than one group, as was the case for A.H., L.P., L.K., R.O., and M.C., three mean treatment quality scores would be calculated (one for each instructional group).

<table>
<thead>
<tr>
<th>Instructor</th>
<th>School</th>
<th>Mean Treatment Quality (scale: 0-21)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Group 1</td>
</tr>
<tr>
<td>D.C.</td>
<td>1</td>
<td>9.4</td>
</tr>
<tr>
<td>D.M.</td>
<td>1</td>
<td>15.0</td>
</tr>
<tr>
<td>A.H.</td>
<td>2</td>
<td>15.0</td>
</tr>
<tr>
<td>L.P.</td>
<td>2</td>
<td>7.3</td>
</tr>
<tr>
<td>L.K.</td>
<td>2</td>
<td>18.0</td>
</tr>
<tr>
<td>R.O.</td>
<td>3</td>
<td>10.6</td>
</tr>
<tr>
<td>S.W.</td>
<td>3</td>
<td>7.9</td>
</tr>
<tr>
<td>M.C.</td>
<td>3</td>
<td>14.6</td>
</tr>
</tbody>
</table>

*Data were supplied for each group that received instruction*

Table 3.3: Mean treatment quality scores by group

_Treatment Duration_

The National Reading Panel’s (2000) meta-analysis on phonemic awareness interventions and optimal length, or duration, of intervention suggests that between 5 and 15 hours of instruction is most beneficial. More or less time than that seems to be inconsequential or detrimental to reading outcomes. It is intuitive that a longer duration of intervention might produce improved outcomes. Due to substantial variability in
intervention implementation consistency and differences in intervention schedules among the schools, a log of time spent in instruction was collected. Establishing a daily log and measuring treatment duration in total minutes is a more sensitive measure than simply reporting that the intervention was implemented for 16 weeks. Within that time frame, there was considerable variability in the consistency and duration of treatment, so it is possible that it may affect the treatment outcomes.

A timer and log sheet was provided for each IA and GA. They were expected to time the instruction and fill in the total instruction time (in minutes and seconds) immediately upon completion of the lesson. Seconds were rounded to the nearest minute. These log sheets were collected weekly or monthly (for School 2). The sum of minutes in instruction was recorded for each group the IA or GA led. Therefore, the exact amount of instruction each student received was the variable of interest. The time log that each instructor filled out is located in Appendix J.

**General Procedures**

The general procedures in this intervention were the same for each IA at each school. Differences were mainly the schedule and consistency of implementation, which was recorded in the daily time log. The procedures for the training session and the intervention in general are discussed here.

**Training**

All Kindergarten classroom teachers, IAs, Reading First coordinators, and principals were invited to participate in the training session, however, training was only required for the IAs. The rationale for including the classroom teachers was to promote acceptance of priority for the intervention. It was hypothesized that if teachers believed
the intervention would be beneficial for students, the intervention would be a priority and not be usurped by clerical or auxiliary activities, such as copying papers or taking students to the restroom. Only the IAs took the training at School 1. All Kindergarten teachers, the data manager, and the IAs at School 2 participated. Only the IAs took the training at School 3.

One hour was allotted for training. The training for Schools 1 and 2 was conducted in one session. School 3 requested that the training be conducted as two 30-minute sessions. After a brief introductory explanation, a seven-minute portion of the demonstration video was shown to give the participants a picture of how the intervention would be conducted. Next, the components and materials of the program were described, with special attention to materials and activities from Part 1. The nature of a scripted curriculum was discussed and a rationale for following the script and adhering to the protocol was provided. Next, the training focused on elements of effective instructional delivery that aren’t unique to the ERI, but are critical for its successful implementation. The definition and examples of active student responding (versus on-task behavior) were given because the ERI requires a high rate of active student responding. Delivery of effective feedback, both reinforcement of correct responses and correction of errors, was demonstrated and discussed. Participants also learned about how to pace instruction and deliver praise. Examples and non-examples of these elements of effective instructional delivery were demonstrated and participants were asked to identify appropriate examples and give suggestions for correcting the non-examples. Finally, the primary researcher modeled several activities from lesson one. Participants then immediately practiced the same activities with a partner. The primary researcher provided feedback and assistance
during this practice session. If time permitted, the pairs were encouraged to practice additional activities. Questions were answered at the conclusion of the session and all participants were instructed to view the demonstration video in its entirety at home.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Objective</th>
<th>Approximate time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>Introduce the participants to the primary researcher and the purpose of the project.</td>
<td>~3 min.</td>
</tr>
<tr>
<td>Video</td>
<td>Provide participants with a visual depiction of the target teaching behavior.</td>
<td>~7 min.</td>
</tr>
<tr>
<td>Program materials</td>
<td>Describe program materials and components. Participants were shown how to identify program materials.</td>
<td>~3 min.</td>
</tr>
<tr>
<td>Effective instructional delivery</td>
<td>Teach participants about instructional delivery techniques that are not in the script such as pacing, praise, feedback and error correction.</td>
<td>~12 min.</td>
</tr>
<tr>
<td>Guided practice/feedback</td>
<td>Practice delivering instruction with guidance and feedback.</td>
<td>~20 min.</td>
</tr>
<tr>
<td></td>
<td>Practice delivering instruction with guidance and feedback.</td>
<td>~20 min.</td>
</tr>
<tr>
<td>Questions</td>
<td>Respond to participants’ questions</td>
<td>~5 min.</td>
</tr>
</tbody>
</table>

Table 3.4: ERI training schedule outline

*Placement Test and Intervention Procedures*

Once students eligible for participation were identified, each student was given the ERI placement test from the assessment book that comes with the curriculum. The placement test is composed of six parts. The first part evaluates the student’s knowledge of letter names and sounds. The rest of the test assesses various phonological awareness and alphabetic principle skills. The results of the placement test determine where the
student needs to start in the curriculum. All of the students in the study began at lesson 1, with the exception of D.C.’s group and one of R.O.’s groups. These students began at the next entry point, lesson 43. When IAs instructed multiple small groups (e.g., School 2) the primary researcher grouped the students according to similar pretest scores on the WJ – III and lower ability groups were assigned to IAs who had more experience and/or demonstrated a higher level of mastery during the training session. For example, the lower ability group at School 1 was assigned to D.M. because she had experience teaching ERI during the previous school year. The decision to pair lower ability students with a more experienced IA was intended to increase the likelihood that the neediest students would receive the best quality instruction possible.

Each day the intervention was scheduled to take place, the IA picked students up from the classroom and took the group to the designated intervention area. Once seated, the IA filled in the daily log sheet, which identified the students in the group, the lesson being taught, whether or not it was a re-teach, and started the timer. Instruction took place for 25-30 minutes and usually one lesson was completed at each session. On rare occasions a lesson had to be carried over to the next session and/or the IA needed to re-teach the lesson if students did not respond appropriately the first time. When the session was over, the IA took the students back to their respective classrooms and returned to their other duties.

Materials

1. Stopwatch. A stopwatch was set for 10 minutes for each observation. The primary research began the observation once the stopwatch was started and concluded the observation when 10 minutes had passed. If an interruption
occurred (e.g. instruction related question, school announcements, etc.), the
timer and the observation were paused.

2. **Timer.** Each IA and GA was provided a small kitchen timer, which could
count up and down in hours, minutes, and seconds. The instructor pressed
start at the beginning of the lesson, and stopped the timer at the conclusion of
the lesson or intervention period.

3. **Folders.** Thirteen standard 8 ½ by 11” folders were used to hold copies of the
materials used in the training sessions.

4. **ERI curriculum.** Six complete ERI curriculum kits were used by the
instructors in this study. Because two schools were Reading First schools, ERI
kits were already purchased for these schools. However, only two of the kits
used in this study were purchased with Reading First funds.

5. **Star card and stamper.** Each student had a grid of 15 boxes that the instructor
stamped with a colorful marker or stamp when the student responded or
behaved appropriately. Once each box was filled the student was eligible to
select a small tangible reward.

6. **Tangible rewards.** Several small tangible rewards were purchased for the
purpose of rewarding students for appropriate behavior and responding. The
tangibles included items such as stickers, pencils, erasers, small balls and toys,
and candy.
Experimental Design

This correlational study used a pretest posttest nonequivalent control group design was used since random selection or assignment was not possible. Treatment students were required to receive the intervention due to risk status as determined by the pretest scores. All students who qualified for treatment received treatment. Control students, then, were not at risk for reading failure, and therefore, not equivalent to the treatment group. The lack of randomization precludes a causal analysis, so correlational analyses were used to explain the relationship between the independent and dependent variables. In the figure below $X_1$ indicates the treatment students who were non-ELLS and $X_2$ indicates the treatment students who were ELLs.

Data Analysis

The research questions in this study evaluate the effects of participation in ERI, quality of implementation, and treatment duration on LWID and WA outcome scores for at risk ELL and non-ELL Kindergarten students. Additionally, the effects of the independent variables on DIBELS intervention recommendation status was investigated. The multivariate techniques multiple linear regression and logistic regression were used to answer these questions.

*Multiple Linear Regression*

Multiple regression analysis was selected to analyze the WJ – III outcome data in this study. Multiple regression analysis is a form of general linear modeling in which
multiple independent variables are regressed on a singular dependent variable (Hair, Anderson, Tatham, & Black, 1998). The purpose of multiple regression analyses is to explain the variance observed in the dependent variable through the linear relationship of the independent variables. The regressions equation is as follows:

\[ Y' = a + b_1X_1 + b_2X_2 + \ldots + b_kX_k \]

Where:

\( Y' \) = Estimate of the dependent variable
\( X_1, X_2, X_k \) = Metric independent variables
\( a \) = estimate of \( Y \) when each independent variable is zero (intercept)
\( b_k \) = partial regression coefficient

1. When other independent variables are held constant, \( b_k \) indicates the expected change in \( Y \) for every one unit change in \( X_k \)
2. The magnitude of the partial regression coefficient is dependent on the unit of measurement used for the independent variable.

The regression equation results in a coefficient of determination (\( R^2 \)). This coefficient estimates the proportion of variance in the dependent variables that can be accounted for by the combination of the independent variables. The interpretation of this analysis is moderated by the following assumptions:

1. No specification errors are present. Additionally, no relevant variables have been excluded and no irrelevant variables have been included. The selection of relevant variables for investigation is based on the theoretical knowledge base of the topic.
2. No measurement error is present. This assumption addresses the accuracy of the measurement of the dependent variables.
3. No systematic errors threaten the interpretation of the data. Errors should be independent. That is, errors associated with any observation are not correlated with errors of other observations or any of the independent variables. Any error in the analysis should be random error. Errors should have a mean of zero, be normally distributed, and have homoscedasticity (constant variance).

4. Perfect collinearity is not present. Collinearity refers to the extent to which independent variables are related to one or more of the other variables in the equation. Perfect collinearity indicates a perfect (1.0) correlation between or among variables.

5. An examination of the residuals indicated that there were no violations.

**Multiple Regression Statistical Models**

1. **Multiple Linear Regressions to model Woodcock Johnson gain scores for control students vs. treatments students, by ELL**

   \[ Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_p X_{pi} + \varepsilon_i ; \quad \text{for } i = 1, 2, \ldots, 96 \]

   Where the error terms \( \varepsilon_i \) are independent, and identically distributed as normal random variables with mean 0 and variance \( \sigma^2 \).

   **Model 1a: Effects of ERI on Word Attack Gain**

   - \( Y_i \): Spring Word Attack
   - \( \beta_0 \): Intercept term
   - \( \beta_p \): Coefficient for \( X_p \)
   - \( X_p \): Predictor (independent) variable: ERI

   **Model 1b: Effects of ERI on Letter Word ID Gain**

   - \( Y_i \): Spring Letter Word ID
   - \( \beta_0 \): Intercept term
   - \( \beta_p \): Coefficient for \( X_p \)
   - \( X_p \): Predictor (independent) variable: ERI
2. Multiple Linear Regressions to model Woodcock Johnson gain scores for ERI students only, by ELL

Model 2a: Effects of TIME and INTEGRITY on Word Attack gain scores within the program

\[ Y_i: \text{Spring Word Attack} \]
\[ \beta_0: \text{Intercept term} \]
\[ \beta_p: \text{Coefficient for } X_p \]
\[ X_p: \text{Predictor (independent) variables: TIME, and INTEGRITY} \]

Model 2b: Effects of TIME and INTEGRITY on Letter Word ID scores within the program

\[ Y_i: \text{Spring Letter Word ID} \]
\[ \beta_0: \text{Intercept term} \]
\[ \beta_p: \text{Coefficient for } X_p \]
\[ X_p: \text{Predictor (independent) variables: TIME, and INTEGRITY} \]

*Effect Size*

A method for describing the magnitude, or the practical significance, of an effect is to calculate an Effect Size (ES). Common effect size calculation methods include Glass’s delta, Hedge’s g, and Cohen’s d. Hedge’s g is an inferential measure and is derived from the square root of the mean square error. Glass’s delta is descriptive in nature, but uses the standard deviation of the control group, so given that the groups are not homogeneous, the standard deviation of the control group may not be reflective of the standard deviation of the experimental group. Therefore, Cohen’s d was selected since it uses the pooled standard deviation of both groups. The pooled standard deviation is the square root of the average of the squared standard deviations.
The magnitude of the treatment still requires interpretation. That is, how robust must the effect be in order to be considered important. Although behavioral science is a broadly diverse field in terms of power analysis, some guidelines can be offered for interpretation. Cohen (1988) indicates that an ES of .2 may be considered small. An ES of .5 is moderate, and .8 and higher is large. The ES can be translated into both average percentile ranking and percentage of non-overlap between treatment and control groups. For example, an ES of .8 indicates that the mean of the treatment group is at the 79th percentile of the control group. An ES of 0.0 would indicate that the mean of the treatment group is at the 50th percentile of the control group. Likewise, an ES of 0.0 would indicate that there is complete overlap in the distribution of treatment and control group scores. In this study effect sizes of participation in the ERI program were calculated for the ELL and non-ELL students separately.

Social Validity

In applied research it is important to evaluate the acceptability of interventions. Three areas should be assessed for acceptance: the target behavior, the procedures used to change that behavior, and the importance of the results obtained (Wolf, 1978). Furthermore, various consumers’ opinions should be collected. Teachers’ and IAs’ opinions of an intervention can determine whether or not the intervention is continued in the classroom. Students’ opinions are important since their level of enthusiasm about a particular intervention can influence how engaged they are in instruction. Finally, the reactions of school administrators and auxiliary personnel may affect whether or not or how the intervention is sustained at the school level.
At the beginning of the study, a pre-acceptability questionnaire was given to all relevant consumers, except students (Appendix K). The questionnaire contained nine items with a five-point agreement scale, ranging from strongly disagree to strongly agree. The questionnaire evaluated the consumer’s perceived anticipated effectiveness of the intervention and procedures. A space was also provided to write in comments. The same questionnaire was again given to each consumer at the conclusion of the study. A semi-structured interview was conducted students as well. The interview form is adapted from the *Children’s Social Validity Interview* (CSVI; Lane, 1997; Appendix L). It is designed to capture the student’s perception of the intervention, the skills acquired, and whether or not those skills are used in other areas.
CHAPTER 4

RESULTS

This chapter presents the results of the study. Specifically, this chapter addresses how much variance in outcome scores can be explained by participation in the early reading intervention, treatment quality, and treatment duration for both ELL and non-ELL students. Additionally, the results of the social validity assessments will be described. The results will be presented in the following sections: (a) Woodcock-Johnson mean scores, and the simple and multiple regression models including mean scores, standard deviations, $R^2$ values, Cohen’s $d$, partial regression coefficients (B), standard error, and Beta values, (b) DIBELS recommendation status, and (c) General results, including the descriptive results of the treatment quality and social validity assessments.

Woodcock-Johnson Tests of Achievement

In order to control for the differences in fall scores, participation in the treatment, treatment duration, and treatment quality were regressed on the gain scores for both the WA and LWID subtests, rather than the raw scores. The mean pretest and posttest scores are described first, followed by the linear regression models. Each of these models was conducted for ELL and non-ELL treatment students separately. Models for WA are presented first, followed by the LWID models.
Mean scores

*Word attack.* The mean pretest score for non-ELL treatment students was 2.1, and the mean posttest score was 5.8. The mean pretest score for ELL treatment students was 1.6, and the mean posttest score was 6.3. The mean pretest score for control students was 3.2, and the mean posttest score was 6.2. The slopes of these gains are graphically depicted in Figure 4.1.

![Figure 4.1: Slope of mean pretest and posttest word attack scores for non-ELL treatment, ELL treatment, and control students.](image)

*Letter word identification.* The mean pretest score for non-ELL treatment students was 9.6, and the mean posttest score was 19.5. The mean pretest score for ELL treatment students was 6.4, and the mean posttest score was 18.5. The mean pretest score for control students was 13.9, and the mean posttest score was 21.3. The slopes of these gains are graphically depicted in Figure 4.2.
Word Attack

Model 1a: Effects of ERI on word attack gain (non-ELL). This model examined the amount of variance in gain scores on the WJ-III word attack subtest that could be attributed to participation in the intervention program for non-ELL students at risk for reading failure. This model used scores from 70 students (38 non-ELL treatment and 32 non-ELL control). As seen in Table 4.1, the total mean gain score was 3.15 with a standard deviation of 1.44. The R value was .381 and the R² value for this model was .15, indicating that 15% of the variance in word attack gain scores could be attributed to participation in the intervention. The effect size was large (.81). The partial regression coefficient (B) was 1.1, indicating that students in the ERI program gained 1.1 points.
more than control students. The standard error was .322 and the standardized Beta coefficient was .381.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>R</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-ELL</td>
<td>70</td>
<td>.381</td>
<td>3.15</td>
<td>1.44</td>
<td>.81</td>
</tr>
<tr>
<td>ELL</td>
<td>55</td>
<td>.561</td>
<td>3.44</td>
<td>1.85</td>
<td>1.32</td>
</tr>
</tbody>
</table>

Table 4.1: Descriptive data for the effects of participation in ERI on word attack gain scores for non-ELL and ELL treatment students.

<table>
<thead>
<tr>
<th></th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>1.1</td>
<td>.322</td>
</tr>
<tr>
<td>ELL</td>
<td>3.44</td>
<td>1.85</td>
</tr>
</tbody>
</table>

R² = .15 (non-ELL), .32 (ELL)

Table 4.2: Coefficients for the effects of participation in ERI on word attack gain scores for non-ELL and ELL treatment students.

Model 1a: Effects of ERI on word attack gain (ELL). This model examined the amount of variance in gain scores on the WJ-III word attack subtest that could be attributed to participation in the intervention program for ELL students at risk for reading failure. This model used scores from 55 students (23 ELL treatment and 32 non-ELL control). The total mean gain score was 3.44 with a standard deviation of 1.85. The R value was .561 and the R² value for this model was .32, indicating that 32% of the variance in word attack gain scores could be attributed to participation in the intervention. The effect size was large (1.32). The partial regression coefficient (B) was 2.1, indicating that students in the ERI program gained 2.1 points more than control students. The standard error was .423 and the standardized Beta coefficient was .561.
Model 2a: Effects of treatment duration and quality on word attack gain (non-ELL). This model examined the amount of variance in gain scores on the WJ-III word attack subtest that could be attributed to the duration of the treatment and the quality with which it was implemented. This model used scores from 38 students (non-ELL treatment only). As indicated in Table 4.3, the mean duration was 595 minutes, which is 9.91 hours of instruction, with a standard deviation of 208.39. The mean quality score was 11.5 on a scale of 21 with a standard deviation of 3.38. The R value was .175 and the $R^2$ value for this model was .03, indicating that 3% of the variance in word attack gain scores could be attributed to treatment duration and quality. The partial regression coefficients for duration and quality were -.001 and .067 indicating that for each unit increase in duration and quality, word attack gain score decreased by .001 and increased by .067, respectively. The standard errors for duration and quality were .001 and .066, respectively. The standardized Beta coefficients were -.094 and .175.

<table>
<thead>
<tr>
<th>Treatment Duration</th>
<th>Treatment Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>R</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>70</td>
</tr>
<tr>
<td>ELL</td>
<td>55</td>
</tr>
</tbody>
</table>

Table 4.3: Descriptive data for the effects of treatment duration and treatment quality on word attack gain scores for non-ELL and ELL treatment students.
<table>
<thead>
<tr>
<th></th>
<th>Treatment Duration</th>
<th>Treatment Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unstandardized</td>
<td>Standardized</td>
</tr>
<tr>
<td></td>
<td>Coefficients</td>
<td>Coefficients</td>
</tr>
<tr>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>-.001</td>
<td>.001</td>
</tr>
<tr>
<td>ELL</td>
<td>-.003</td>
<td>.002</td>
</tr>
</tbody>
</table>

$R^2 = .03$ (non-ELL), $.39$ (ELL)

Table 4.4: Coefficients for the effects of treatment duration and treatment quality on word attack gain scores for non-ELL and ELL treatment students.

**Model 2a: Effects of treatment duration and quality on word attack gain (ELL).**

This model examined the amount of variance in gain scores on the WJ-III word attack subtest that could be attributed to the duration of the treatment and the quality with which it was implemented. This model used scores from 23 students (ELL treatment only). The mean duration was 289 minutes, which is 4.82 hours of instruction, with a standard deviation of 129.74. The mean quality score was 16.8 on a scale of 21 with a standard deviation of 1.81. The R value was .621 and the $R^2$ value for this model was .39, indicating that 39% of the variance in word attack gain scores could be attributed to treatment duration and quality. The partial regression coefficients for duration and quality were -.003 and -.491, indicating that for each unit increase in duration and quality, word attack gain score decreased by .003 and .491, respectively. The standard errors for duration and quality were .002 and .170, respectively. The standardized Beta coefficients were -.321 and -.508.

**Letter Word Identification**

**Model 1b: Effects of ERI on letter word identification gain (non-ELL).** This model examined the amount of variance in gain scores on the WJ-II letter word identification subtest that could be attributed to participation in the intervention program for non-ELL
students at risk for reading failure. This model used scores from 70 students (38 non-ELL treatment and 32 non-ELL control). Table 4.5 indicates that the total mean gain score was 9.20 with a standard deviation of 3.66. The R value was .335 and the R² value for this model was .11, indicating that 11% of the variance in letter word identification gain scores could be attributed to participation in the intervention. The effect size was large (.99). The partial regression coefficient (B) was 2.4, indicating that students in the ERI program gained 2.4 points more than control students. The standard error was .833 and the standardized Beta coefficient was .335.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>R</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-ELL</td>
<td>70</td>
<td>.335</td>
<td>9.20</td>
<td>3.66</td>
<td>.99</td>
</tr>
<tr>
<td>ELL</td>
<td>55</td>
<td>.458</td>
<td>9.60</td>
<td>4.49</td>
<td>.70</td>
</tr>
</tbody>
</table>

Table 4.5: Descriptive data for the effects of participation in ERI on letter word identification gain scores for non-ELL and ELL treatment students.

<table>
<thead>
<tr>
<th></th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>2.4</td>
<td>.833</td>
</tr>
<tr>
<td>ELL</td>
<td>4.1</td>
<td>1.10</td>
</tr>
</tbody>
</table>

R² = .11 (non-ELL), .21 (ELL)

Table 4.6: Coefficients for the effects of participation in ERI on letter word identification scores for non-ELL and ELL treatment students.

Model 1b: Effects of ERI on letter word identification gain (ELL). This model examined the amount of variance in gain scores on the WJ-III letter word identification subtest that could be attributed to participation in the intervention program for ELL students at risk for reading failure. This model used scores from 55 students (23 ELL
treatment and 32 non-ELL control). The total mean gain score was 9.60 with a standard deviation of 4.49. The R value was .458 and the $R^2$ value for this model was .21, indicating that 21% of the variance in letter word identification gain scores could be attributed to participation in the intervention. The effect size was moderate (.70). The partial regression coefficient (B) was 4.1, indicating that students in the ERI program gained 4.1 points more than control students. The standard error was 1.10 and the standardized Beta coefficient was .458.

Model 2b: Effects of treatment duration and quality on letter word identification gain (non-ELL). This model examined the amount of variance in gain scores on the WJ-III letter word identification subtest that could be attributed to the duration of the treatment and the quality with which it was implemented. This model used scores from 38 students (non-ELL treatment only). As seen in Table 4.7, the mean duration was 595 minutes, which is 9.91 hours of instruction, with a standard deviation of 208.39. The mean quality score was 11.5 on a scale of 21 with a standard deviation of 3.38. The R value was .473 and the $R^2$ value for this model was .22, indicating that 22% of the variance in letter word identification scores could be attributed to treatment duration and quality. The partial regression coefficients for duration and quality were .008 and -.160, indicating that for each unit increase in duration and quality, letter word identification gain score increased by .008 and decreased by .160, respectively. The standard errors for duration and quality were .003 and .161, respectively. The standardized Beta coefficients were .491 and -.154.
<table>
<thead>
<tr>
<th>Treatment Duration</th>
<th>Treatment Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>70</td>
</tr>
<tr>
<td>ELL</td>
<td>55</td>
</tr>
</tbody>
</table>

Table 4.7: Descriptive data for the effects of treatment duration and treatment quality on letter word identification gain scores for non-ELL and ELL treatment students.

<table>
<thead>
<tr>
<th>Treatment Duration</th>
<th>Treatment Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unstandardized Coefficients</td>
</tr>
<tr>
<td></td>
<td>B</td>
</tr>
<tr>
<td>Non-ELL</td>
<td>.008</td>
</tr>
<tr>
<td>ELL</td>
<td>-.004</td>
</tr>
</tbody>
</table>

R² = .22 (non-ELL), .26 (ELL)

Table 4.8: Coefficients for the effects of treatment duration and treatment quality on letter word identification gain scores for non-ELL and ELL treatment students.

Model 2b: Effects of treatment duration and quality on letter word identification gain (ELL). This model examined the amount of variance in gain scores on the WJ-III letter word identification subtest that could be attributed to the duration of the treatment and the quality with which it was implemented. This model used scores from 23 students (ELL treatment only). The mean duration was 289 minutes, which is 4.82 hours of instruction, with a standard deviation of 129.74. The mean quality score was 16.8 on a scale of 21 with a standard deviation of 1.81. The R value was .507 and the R² value for this model was .26, indicating that 26% of the variance in letter word identification scores could be attributed to treatment duration and quality. The partial regression coefficients for duration and quality were -.004 and –1.219, indicating that for each unit increase in duration and quality, letter word identification gain score decreased by .004 and 1.219,
respectively. The standard errors for duration and quality were .005 and .509, respectively. The standardized Beta coefficients were -.174 and -.463.

DIBELS Benchmark Assessments

The results of the spring benchmark assessments for each group will be described along with the number and percentage of students who improved their recommendation status, stayed the same, or regressed.

Non-ELL Treatment Students

In the fall, 34.2% (13) needed intensive intervention, 65.7% (25) were strategic, and none of the non-ELL treatment students were at benchmark. In the spring, 28.9% (11) of the students needed intensive intervention, 28.9% (11) received a strategic intervention recommendation, and 42.1% (16) of the students were at benchmark. See Figures 4.3 to compare proportional recommendations in fall and spring. This change represents an improvement for 50% (19) of the students, 42.1% (16) remained the same, and 7.8% (3) regressed.
Figure 4.3: Proportion of non-ELL treatment students with intensive, strategic, and benchmark intervention recommendations in fall and spring.

**ELL Treatment Students**

In the fall, 43.4% (10) needed intensive intervention, 47.8% (11) were strategic, and 8.6% (2) were benchmark. In the spring, 43.4% (10) of the students needed intensive intervention, 17.4% (4) received a strategic intervention recommendation, and 39.1% (9) of the students were at benchmark. This change represents an improvement for 43.4% (10) of the students and 56.5% (13) remained the same. None regressed.
Figure 4.4: Proportion of ELL treatment students with intensive, strategic, and benchmark intervention recommendations in fall and spring.

Control Students

In the fall, none of the students needed intensive intervention, 12.5% (4) were strategic, and the remaining 87.5% (28) students were at benchmark. In the spring, 9.3% (3) of the students needed of intensive intervention, 25% (8) received a strategic intervention recommendation, and 65.6% (21) of the students were at benchmark. This change represents an improvement for 3.1% (1) of the students, 68.7% (22) remained the same, and 28.1% (9) regressed.
Figure 4.5: Proportion of control students with intensive, strategic, and benchmark intervention recommendations in fall and spring.

**General Results**

This section will describe the results of the direct observations of treatment quality and include total treatment duration and students’ mean percentage gain on LWID and WA for each individual instructor. Additionally, the results of the social validity questionnaires for the school principals, related school personnel, instructors, and students will be described.

*Treatment Quality*

The seven components of treatment quality were scored on a scale of 0 to 21 for each observation session. The treatment quality observation form is located in Appendix I. Due to instructor absences or school-related functions, not all instructors have the same number of data points, although every attempt was made to make up missed observations.

*D.C. (School 1)*. A total of 12 observation sessions yielded a treatment quality range between 6 (28.5%) and 15 (71.4%) points. D.C.’s mean treatment quality was 9.4 (44.7%). Total treatment duration was 6.85 hours of instruction. The students in D.C.’s
group had a mean pretest score of 13.6 (range 8-17) raw score points on LWID and a mean posttest score of 21.6 (range 14-25), resulting in a mean gain of 8.0, which represents 37% of the posttest score. On WA, D.C.’s group had a mean pretest score of 2.6 (range 1-3) raw score points and a mean posttest score of 6.2 (range 5-7), resulting in a mean gain of 3.6, which represents 58% of the posttest score.

Figure 4.6: Treatment quality scores for D.C. with a mean line.

**D.M. (School 1).** A total of 9 observation sessions yielded a treatment quality range between 10 (47.6%) and 18 (85.7%) points. D.M.’s mean treatment quality was 15.0 (71.4%). Total treatment duration was 7.55 hours of instruction. The students in D.M.’s group had a mean pretest score of 10.0 (range 8-17) raw score points on LWID and a mean posttest score of 17.0 (range 3-18), resulting in a mean gain of 7.0, which represents 41% of the posttest score. On WA, D.M.’s group had a mean pretest score of 1.2 (range 0-3) raw score points and a mean posttest score of 5.0 (range 3-7), resulting in a mean gain of 4.8, which represents 96% of the posttest score.
Figure 4.7: Treatment quality scores for D.M. with a mean line.

For visual interpretation of these data, Figure 4.8 shows the mean percentage treatment quality for D.C. and D.M. juxtaposed upon mean gains for their students. Total treatment duration is also given. These data from IAs at School 1 suggest a possible relationship between treatment quality and duration and students gains. In this case, D.M.’s students made greater gains than D.C.’s students. D.M. also had a greater duration of treatment as well as higher quality of implementation.
Figure 4.8: School 1 instructors’ mean percentage treatment quality, duration, and student gains on LWID and WA posttest.

*A.H.* *(School 2).* A total of 16 observation sessions yielded a treatment quality range between 9 (42.8%) and 19 (90.4%) points. A.H.’s mean treatment quality was 15.6 (74.2%). Total mean treatment duration for his three groups was 13.70 hours of instruction (range 13.13-14.66). Groups 1, 2, and 3 are indicated on Figure 4.9 with different data point shapes. The students in A.H.’s groups had a mean pretest score of 8.1 (range 5-13) raw score points on LWID and a mean posttest score of 20.2 (range 16-25), resulting in a mean gain of 12.1, which represents 60% of the posttest score. On WA, A.H.’s groups had a mean pretest score of 2.3 (range 1-3) raw score points and a mean posttest score of 6.2 (range 4-7), resulting in a mean gain of 3.9, which represents 63% of the posttest score.
A total of 16 observation sessions yielded a treatment quality range between 3 (14.2%) and 12 (57.1%) points. L.P.’s mean treatment quality was 8.1 (38.5%). Total mean treatment duration for her three groups was 12.84 hours of instruction (range 12.66-12.96). Groups 1, 2, and 3 are indicated on Figure 4.9 with different data point shapes. The students in L.P.’s groups had a mean pretest score of 8.6 (range 3-13) raw score points on LWID and a mean posttest score of 19.9 (range 11-26), resulting in a mean gain of 11.3, which represents 57% of the posttest score. On WA, L.P.’s groups had a mean pretest score of 2.6 (range 1-4) raw score points and a mean posttest score of 6.1 (range 4-11), resulting in a mean gain of 3.5, which represents 57% of the posttest score.
Figure 4.10: Treatment quality scores for L.P. with a mean line.

L.K. (School 2). A total of 18 observation sessions yielded a treatment quality range between 17 (80.9%) and 20 (95.2%) points. L.K.’s mean treatment quality was 18.4 (87.6%). Total mean treatment duration for her three groups was 5.76 hours of instruction (range 2.38-10.90). Groups 1, 2, and 3 are indicated on Figure 4.9 with different data point shapes. The students in L.K.’s groups had a mean pretest score of 9.4 (range 0-22) raw score points on LWID and a mean posttest score of 19.1 (range 9-29), resulting in a mean gain of 9.7, which represents 51% of the posttest score. On WA, L.K.’s groups had a mean pretest score of 2.1 (range 0-7) raw score points and a mean posttest score of 5.8 (range 1-12), resulting in a mean gain of 3.7, which represents 64% of the posttest score.
Figure 4.11: Treatment quality scores for L.K. with a mean line.

For visual interpretation of these data, Figure 4.12 shows the mean percentage treatment quality for A.H., L.P., and L.K. juxtaposed upon mean gains for their students. Total treatment duration is also given. These data from IAs at School 2 suggest a possible relationship between treatment quality and duration and students gains for the non-ELL students. In this case, A.H.’s students made greater gains than L.P’s students. A.H. also had a greater duration of treatment as well as higher quality of implementation. Although L.K.’s students had shorter treatment duration, this deficit may have been mitigated by an exceptionally high level of treatment quality, resulting in considerable gains, particularly for word attack.
R.O. *(School 3).* A total of 15 observation sessions yielded a treatment quality range between 7 (33.3%) and 15 (71.4%) points. R.O.’s mean treatment quality was 11.2 (53.3%). Total mean treatment duration for her three groups was 8.09 hours of instruction (range 6.36-9.83). Groups 1 and 2 are indicated on Figure 4.9 with different data point shapes. The students in R.O.’s groups had a mean pretest score of 8.1 (range 4-12) raw score points on LWID and a mean posttest score of 17.9 (range 14-21), resulting in a mean gain of 9.8, which represents 55% of the posttest score. On WA, R.O.’s groups had a mean pretest score of 2.1 (range 1-3) raw score points and a mean posttest score of 5.2 (range 3-7), resulting in a mean gain of 3.1, which represents 60% of the posttest score.
Figure 4.13: Treatment quality scores for R.O. with a mean line.

*S.W. (School 3).* A total of 15 observation sessions yielded a treatment quality range between 4 (19.0%) and 11 (52.3%) points. S.W.’s mean treatment quality was 7.9 (37.6%). Total mean treatment duration for her three groups was 6.67 hours of instruction (range 6.45-6.90). Groups 1 and 2 are indicated on Figure 4.9 with different data point shapes. The students in S.W.’s groups had a mean pretest score of 8.6 (range 5-12) raw score points on LWID and a mean posttest score of 19.0 (range 17-21), resulting in a mean gain of 10.4, which represents 55% of the posttest score. On WA, S.W.’s groups had a mean pretest score of 1.1 (range 1-2) raw score points and a mean posttest score of 5.7 (range 3-7), resulting in a mean gain of 4.6, which represents 81% of the posttest score.
Figure 4.14: Treatment quality scores for S.W. with a mean line.

*M.C. (School 3).* A total of 16 observation sessions yielded a treatment quality range between 9 (42.8%) and 20 (95.2%) points. M.C.’s mean treatment quality was 15.3 (72.8%). Total mean treatment duration for her three groups was 3.59 hours of instruction (range 1.60-6.23). Groups 1, 2, and 3 are indicated on Figure 4.9 with different data point shapes. The students in M.C.’s groups had a mean pretest score of 2.9 (range 0-10) raw score points on LWID and a mean posttest score of 17.4 (range 14-21), resulting in a mean gain of 14.5, which represents 83% of the posttest score. On WA, M.C.’s groups had a mean pretest score of 0.8 (range 0-2) raw score points and a mean posttest score of 6.5 (range 5-7), resulting in a mean gain of 5.7, which presents 88% of the posttest score.
Figure 4.15: Treatment quality scores for M.C. with a mean line.

For visual interpretation of these data, Figure 4.16 shows the mean percentage treatment quality for R.O., S.W., and M.C. juxtaposed upon mean gains for their students. Total treatment duration is also given. These data from IAs at School 3 suggest a possible relationship between treatment quality, but not duration, and students’ gains for the ELL students. In this case, M.C.’s students made greater gains than non-ELL students instructed by R.O. and S.W. However, M.C.’s students received the shorter duration of treatment. As with ELL students from School 2, this again suggests that an exceptionally high level of treatment quality may have mitigated shorter treatment duration. For the non-ELL students, there is no such clear connection between treatment duration and quality and students gains.
Social Validity

The measurement of the direct consumers’ perception of social validity, the extent to which the intervention has changed the participant’s life in socially important ways, is an important aspect in evaluating applied research (Cooper, Heron, & Heward, 1987). It is not only important to establish the effectiveness of interventions, but it’s also important to ensure the usefulness and acceptability. Each instructor, school principal, school-level Reading First coordinator, and data manager was asked to complete the same questionnaire regarding the acceptability of the intervention, its procedures, and the outcomes of the study. The questionnaire contained nine items with a likert-type five-level agreement scale to indicate the extent to which he or she agrees or disagrees with the statement. The scale ranges from 1 – Strongly Disagree to 5 – Strongly Agree. A few lines were included as a space for free response to the intervention. Appendix K shows the complete questionnaire. Student participants were asked for their opinions about the

Figure 4.16: School 3 instructors’ mean percentage treatment quality, duration, and student gains on LWID and WA posttest.
intervention, particularly the extent to which they use the skills learned in the program in other settings. Students were asked to respond to each question by either the primary researcher or L.K., a graduate student in special education. Items were a combination of likert-type scale responses, free response, and yes or no questions. Appendix L shows the complete questionnaire used with the students. Responses were collected from all ELL treatment students. Responses were collected from 32 (84.2%) of the non-ELL treatment students. Six responses were missing due to absences or other school-related interferences. The responses are presented here.

Principal 1. Principal 1 “Strongly Agreed” (5) with every statement on the questionnaire except item 4, which asks about the procedures of the intervention. Principal 1 did not feel she was familiar enough with the specific procedures to make a determination for this statement, so she left the item blank. For item 6, which states that the student will experience discomfort during the intervention, Principal 1 “Strongly Disagreed” (1). Principal 1 did not provide any written comments.

Principal 2. Principal 2 was the only principal to return a completed a pre-intervention acceptability assessment, which was the same questionnaire used for the post-intervention social validity assessment. Other than statement 6, which she rated 1 (strongly disagree), she provided ratings of 4 (agree) and 5 (strongly agree) for all other items. When she completed the social validity questionnaire at the conclusion of the study, the items she rated as 4 on the pre-intervention assessment, she rated as 5, indicating that her perception of the intervention slightly improved by the conclusion of the study. In the space to provide additional comments, Principal 2 offered, “Thank you!”
Principal 3. She rated provided a rating of 3 (neutral) or 4 (agree) for all questionnaire items, except statement 6, indicating that the student will experience discomfort, which she provided a rating of 2 (disagree). Principal 3 did not provide any written comments.

School Personnel. Although the questionnaires were given to the Reading First coordinators and data managers at both Reading First designated schools (Schools 2 and 3), only the Reading First coordinator from School 3 and the data manager from School 2 returned them. Both of these individuals provided positive ratings (4-5) for all items. The data manager disagreed with item 6, but the Reading First coordinator rated this item 5 (strongly agree). Since she provided a rating of 5 for all of the items, it is possible that her rating of item 6 reflects a response set. That is, she completed the questionnaire without carefully reading all items and provided the same rating for every item.
<table>
<thead>
<tr>
<th>Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I find this intervention to be an acceptable way of dealing with students’ reading difficulties</td>
<td>P3</td>
<td>D.M.; RFC; P1; P2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. I would be willing to use this procedure if I had to improve a student’s reading difficulties.</td>
<td>P3</td>
<td>D.M.; RFC; P1; P2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. I believe it would be acceptable to use this intervention without students’ consent.</td>
<td>P3</td>
<td>D.M.; RFC; P1; P2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. I liked the procedures used in this intervention.</td>
<td>P3</td>
<td>D.M. RFC; P2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. I believe this intervention is likely to be effective.</td>
<td>P3</td>
<td>D.M.; RFC; P1; P2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. I believe the students will experience discomfort during the intervention.</td>
<td>P1</td>
<td>D.M.; RFC; P2; P3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. I believe this intervention is likely to result in permanent improvement.</td>
<td>P3</td>
<td>D.M.; RFC; P1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. I believe it would be acceptable to use this intervention with students who cannot choose interventions for themselves.</td>
<td>P3</td>
<td>D.M.; RFC; P1; P2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Overall, I have a positive reaction to this intervention.</td>
<td>D.M.; RFC; P1; P2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.9: Results of the social validity questionnaire for school principals, data manager (School 2) and Reading First coordinator (School 3).

*D.C. (School 1)*. D.C. rated most items a 3 or 4. Item 4, “I like the procedures used in this intervention”, was rated 2 (disagree). Additionally, he provided a neutral rating for item 9, “Overall, I have a positive reaction to this intervention”. He did not provide any additionally written comments.
D.M. (School 1). D.M. rated all items as either 3 (neutral) or 4 (agree). She rated item 6 a 3 (neutral). In the space provided to write comments, D.M. offered, “I would like to see this intervention being used at the beginning of the school year after students have been evaluated by their teachers”.

A.H. (School 2). A.H. returned the pre-intervention acceptability assessment, which was the same form used at post-intervention. Prior to the onset of the study, A.H. gave each item a positive rating of either 4 or 5, except item 6, which he rated 2 (disagree). On the post-intervention questionnaire, He was slightly less positive, rating most items 4. Other ratings were 3. He did not provide an written comments.

L.P. (School 2). On the pre-intervention acceptability assessment, L.P. rated every item 4, except item 6, which she rated 2. Her perception of the intervention improved slightly on the post-intervention assessment, on which she rated every item 5, except item 6, which she again rated 2. She did not provide any written comments.

L.K. (School 2). L.K. rated every item either a 4 or 5. However, for item 6 she provided a rating of 2.

R.O. (School 3). R.O. strongly disagreed (rating=1) that students would experience discomfort during the intervention (item 6). She rated all other items a 3, 4, or 5, with the majority receiving a rating of 4. She did not provide any written comments.

S.W. (School 3). All S.W.’s ratings were between 1 (strongly disagree) and 3 (neutral). She provided a neutral rating for item 6, and strongly disagreed on item 4, indicating that she did not like the procedures used in the intervention. She also strongly disagreed with item 8, “I believe it would be acceptable to use this intervention with
students who can not choose interventions for themselves”, and item 9, “Overall, I have a
good positive reaction to this intervention”. She did not provide any written comments.

*M.C. (School 3).* M.C. rated most items 4 or 5. She strongly disagreed with item
6, and rated items 3 and 8 as neutral. She wrote brief questions next to these items such
as, “They don’t choose (or consent)?” M.C. did not provide any written comments.
<table>
<thead>
<tr>
<th>Item</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I find this intervention to be an acceptable way of dealing with students’ reading difficulties</td>
<td>S.W.</td>
<td>A.H.;</td>
<td>L.P.;</td>
<td>L.K.;</td>
<td>M.C.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D.C.;</td>
<td>D.M.;</td>
<td></td>
<td>R.O.</td>
</tr>
<tr>
<td>2. I would be willing to use this procedure if I had to improve a student’s reading difficulties.</td>
<td>S.W.</td>
<td>A.H.;</td>
<td>L.P.;</td>
<td>L.K.;</td>
<td>M.C.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D.C.;</td>
<td>D.M.;</td>
<td></td>
<td>R.O.</td>
</tr>
<tr>
<td>3. I believe it would be acceptable to use this intervention without students’ consent.</td>
<td>A.H.;</td>
<td>D.C.;</td>
<td>L.P.;</td>
<td>L.K.;</td>
<td>R.O.</td>
</tr>
<tr>
<td></td>
<td>M.C.;</td>
<td>D.M.;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S.W.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. I liked the procedures used in this intervention.</td>
<td>S.W.</td>
<td>D.C.</td>
<td>D.M.;</td>
<td>A.H.;</td>
<td>L.P.</td>
</tr>
<tr>
<td></td>
<td>R.O.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. I believe this intervention is likely to be effective.</td>
<td>S.W.</td>
<td>A.H.;</td>
<td>L.P.;</td>
<td>L.K.;</td>
<td>R.O.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D.C.;</td>
<td>D.M.;</td>
<td></td>
<td>M.C.</td>
</tr>
<tr>
<td>6. I believe the students will experience discomfort during the intervention.</td>
<td>M.C.;</td>
<td>L.P.;</td>
<td>A.H.;</td>
<td>L.K.;</td>
<td>R.O.</td>
</tr>
<tr>
<td></td>
<td>R.O.</td>
<td>D.C.;</td>
<td>D.M.;</td>
<td></td>
<td>S.W.</td>
</tr>
<tr>
<td>7. I believe this intervention is likely to result in permanent improvement.</td>
<td>D.C.;</td>
<td>A.H.;</td>
<td>L.P.;</td>
<td>L.K.;</td>
<td>R.O.</td>
</tr>
<tr>
<td></td>
<td>R.O.;</td>
<td>D.M.;</td>
<td></td>
<td></td>
<td>M.C.</td>
</tr>
<tr>
<td></td>
<td>S.W.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. I believe it would be acceptable to use this intervention with students who cannot choose interventions for themselves.</td>
<td>S.W.</td>
<td>D.M.;</td>
<td>A.H.;</td>
<td>L.P.;</td>
<td>R.O.</td>
</tr>
<tr>
<td></td>
<td>M.C.</td>
<td>D.C.;</td>
<td></td>
<td></td>
<td>L.K.</td>
</tr>
<tr>
<td>9. Overall, I have a positive reaction to this intervention.</td>
<td>S.W.</td>
<td>D.C.;</td>
<td>A.H.;</td>
<td>L.P.;</td>
<td>R.O.</td>
</tr>
<tr>
<td></td>
<td>D.M.;</td>
<td></td>
<td></td>
<td></td>
<td>M.C.</td>
</tr>
</tbody>
</table>

Table 4.10: Results of the social validity questionnaire for the instructors.

*Non-ELL treatment students.* Either L.K. or the primary researcher collected responses for 32 non-ELL treatment students. The format of the items were open-ended, yes or no, or a likert-type preference scale with three levels ranging between “not much”
and “a lot”. A summary of the percentage of responses in each category is included below. Relevant open ended responses included, “I liked learning (sounding out, tracing) the letters (sounds, words)”, “I liked playing games (reading books, earning prizes)”, “I wish the group would last longer (a whole week, 100 days, a couple more hours, 20 more minutes)”, and “I use what I learned at home (school)”. 
Question Students’ Responses (n=32)

Likert-type scale questions.

Did you like:

A. Being part of the group?
B. Spending time with Ms./Mr. ________?
C. Earning rewards/stars?
D. Learning new skills?

Likert-type scale questions.

Did you learn things that will help you do better work:

A. In school
B. At home

Yes/ No Questions.

A. Did you like being in this special group/program?
B. Do you feel you learned important things?
C. Do you use the skills that you learned in the special program/group?
D. Do you wish the group/program could have lasted longer?

Table 4.11: Summary of responses by non-ELL treatment students on social validity questionnaire.
ELL treatment students. Either L.K. or the primary researcher collected responses for 23 ELL treatment students. The format of the items were open-ended, yes or no, or a likert-type preference scale with three levels ranging between “not much” and “a lot”. A summary of the percentage of responses in each category is included below. Relevant open-ended responses included, “I liked learning how to put letters to words (sounds, letters)”, “I liked playing games (earning prizes, stickers)”, “I wish the group would last longer (everyday)”, and “I use what I learned in our class (home, school)”. Often ELL treatment students would respond to items with letter sounds. For example, when asked what was the most important thing learned, several students responded by making the following sounds: /m/, /b/, and /p/. 
Table 4.12: Summary of ELL treatment students’ social validity questionnaire responses.
CHAPTER 5

DISCUSSION

This chapter discusses the results of the study as they relate to the five research questions posed in Chapter 1. The current study examined the effects a supplemental early reading intervention program on urban at-risk students’ phonemic awareness and alphabetic principle skills. Six district-employed instructional assistants and two master’s level graduate assistants delivered the intervention to 61 students. Twenty-three were ELL and 28 were non-ELL. The frequency of instruction was between 3 and 5 times per week, with total treatment duration ranging between 6.85 and 13.70 hours. Data were retained for 32 non-treatment students for comparison purposes. Students were assessed with the DIBELS and the Woodcock-Johnson letter word identification and word attack subtests in the fall and spring. All instructors were observed once or twice per week for level of treatment quality, which encompassed both the level of adherence to the protocol and the extent to which those components were implemented. Participation in the program, treatment duration, and treatment quality were regressed on the dependent variables. Additionally, measures of social validity were taken to evaluate the participants’ opinion of the treatment. This chapter concludes with a summary of the study, limitations, implications for practice, and directions for future research.
Research Question One

How much variance in student outcome scores as measured by the letter-word identification (LWID) and word attack (WA) subtests from the Woodcock-Johnson Tests of Achievement III (WJ III) can be explained by participation in the ERI for both ELL and non-ELL students?

ERI treatment data showed beneficial effects for both non-ELL and ELL students at risk for reading failure, with particularly strong outcomes for the ELL students. Both groups made substantial gains on the WJ-III subtests and compare favorably to their more competent peers. The linear relationship between participation in the program and the outcome scores approached the moderate range for non-ELLs: LWID (.35) and WA (.38) and achieved stronger moderate correlations for the ELL students (i.e., .56 and .46). Additionally, the gain scores on these subtests for both groups were greater than those for controls. As seen in Table 4.1, both treatment groups made more gains on the WJ-III subtests than the controls, with the most robust gains for the ELLs. Considering that the treatment students began the year far behind their control student peers, it is remarkable that the treatment students achieved greater gains than their more competent peers. This is particularly impressive for the ELLs.

The non-ELL model explained 11% and 15% of the variance on LWID and WA, respectively. However, the variance explained for ELLs was substantially greater with 21% and 32% variance explained. It is possible that there is more variance explained for the ELLs due to their limited English proficiency. That is, non-ELL students who have greater capacities for expressive and receptive language in the classroom may benefit more in the regular classroom instruction, thereby diminishing the proportion of unique
variance explained by the ERI program. However, for students who do not speak or understand English as well, it may be difficult to reap those same benefits from regular classroom instruction, and more benefits were gleaned from the explicit, intensive instruction provided in the ERI program.

Still, the effect sizes were large for the most part, lending practical value to the findings. The effect sizes for non-ELLs and ELLs on word attack were .81 and 1.32, respectively. On LWID, the effect sizes for these two groups were .99 and .70. These effect sizes support the National Reading Panel’s (NRP; 2000) conclusions that phonemic awareness instruction results in substantial benefit for students’ phonemic awareness skills. Their meta-analyses found that the overall effect size for phonemic awareness outcomes was .86. In this case, all of the effect sizes approached that level, and the effect size for ELLs on the word attack subtest far surpassed it. Furthermore, the NRP’s meta-analysis on phonics instruction produced an overall effect size of .44, which was well below the effect sizes seen in this study.

The findings of this study demonstrate that some of the unique variance in gain scores on the outcomes measures can be attributed to participation in the ERI program. The amount of variance explained was greater for ELLs, suggesting that while some proportion of the outcomes scores may have been attributed to other environmental variables (e.g. regular classroom instruction, greater receptive and expressive English skills etc.), ELLs may be less likely to “pick it up” from the environment and these students require more direct instruction on early reading and language skills. The results of this study further suggest that early reading interventions conducted in English are not only appropriate for ELLs, it appears that ELLs may reap even more benefit from the
instruction than their English-speaking peers. In this study, master’s level students in special education, L.K. and M.C, taught the ELLs. At School 2, L.K. yielded the highest overall treatment quality scores. M.C. had relatively strong quality scores as well. It is possible that these high quality scores influenced the superior performance of the ELLs.

It should also be noted that although the ELLs had very low pretest scores, their true abilities might not have been apparent due to language differences. The ELLs in this study may have had a wealth of reading potential that was “unlocked” once their English-language skills began to develop.

Overall, these findings are supportive of the literature base, which suggests that early intervention in reading for non-ELLs can improve students’ phonemic awareness skills and possibly prevent future reading difficulties (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Snow, Burns, & Griffin, 1998; Torgesen, Wagner, & Rashotte, 1997b; Vellutino et al., 1996). The results for ELLs demonstrate clear benefits of supplementary reading instruction in English. This is in contrast to primary-language advocates, who recommend that reading instruction first take place in the student’s primary language, and then transfer those skills to English (Thomas & Collier, 1997). Instead, these results support the belief that putting off English-language reading is an unnecessary delay and ELLs can begin to learn in English first (Fitzgerald, & Noblit, 1999; Foorman et al., 2004; Gunn et al., 2005; Lenters, 2004).
Research Question Two

How much variance in student outcome scores as measured by LWID and WA subtests from the WJ III can be explained by treatment duration and treatment quality for both ELL and non-ELL students?

The model explained 22% and 26% of the variance in students’ gain scores on LWID for non-ELLs and ELLs, respectively. For WA gain score, the model explained 3% and 39% of the variance for non-ELLs and ELLs respectively. Except for the WA model for non-ELLs, the unique variance explained by this model is greater than that for the participation in ERI variable. However, the results of this model are mixed.

For non-ELL students, there was a slightly positive linear relationship for treatment duration and LWID gain scores. The partial regression coefficient (B) indicated that for every unit increase in duration (e.g. one minute), gain scores increase by .008. Interestingly, for every unit increase in treatment duration, non-ELL gain scores on WA actually decrease by .001. However, these correlations are miniscule, so the practical meaning may be negligible. The variability in treatment duration was naturally occurring, that is, we did not specifically manipulate this variable. The range of treatment duration was between 6.85 hours and 13.70 hours of instruction. The meta-analyses of the NRP (2000) determined that the optimal range of treatment duration is between 5 and 18 hours of instruction. Treatment durations that fall within this window produce the most optimal outcomes, however individual differences obtained by any durations within this window are still not known. Since the naturally occurring range of treatment durations in this study all fall within this optimal window, these findings support the position taken by the NRP.
Similar negligible correlations were seen for ELLs. On the LWID subtest, for every additional minute of treatment, ELLs gain scores decreased by .003. For WA, their gain scores decreased by .004. One possible explanation for the negative correlation between ELLs gain scores and treatment duration is that these students were part of a multiple baseline design study conducted by the two graduate students in special education. As part of their master’s requirements, this intervention project served as their theses. These studies employed a quasi-experimental multiple baseline design. One of the advantages of this design is that several students needing the same intervention can be accommodated without the withdrawal of treatment. Students were assigned to three groups according to the results of their placement tests. DIBELS progress monitoring data were taken for several weeks to establish baseline responding. Once Group 1 was stable, they began the intervention phase. Once intervention responding was stable for Group 1 and baseline responding for Group 2 was still low and stable, Group 2 entered intervention. Group 3 entered after responding was stable for Group 2. In this design each student serves as his or her own control and the intervention effectiveness is based on the comparison to the student’s baseline levels. However, due to the staggered nature of the design, not all of the students received the same duration of treatment, and overall treatment duration was considerably shorter than their non-ELL peers. Research in special education is concerned with the outcomes of students in the greatest need. Therefore, as an ethical consideration, students who were at the greatest risk and evidenced the greatest deficits entered intervention first. Group 2 entered intervention next and was considered more at risk than Group 3, who entered last. This method for entry to intervention ensures that the neediest students received the longest treatment.
duration, and students who did not present the greatest risk received the shortest duration. This design confounds any conclusion regarding the effects of treatment duration for ELLs.

For non-ELLs this model indicated that for every point increase in quality, students actually gained .160 fewer points on the LWID subtest; however, a positive correlation was seen for WA. For every unit increase in treatment quality, students gained .067 points more. These findings are mixed and difficult to interpret since the coefficients are very small. Many things may have contributed to the ambiguity of the results. The most likely reason may be due to how the instructional assistants were assigned to student groups. Once student groups were formed, the groups that demonstrated the greatest risk were assigned to the IA who demonstrated the most competency with delivery of the intervention, as determined by the primary researcher at the conclusion of each training session. For example, in School 1, D.M. had one year’s experience teaching the program already and had taken several college level education courses. The lower group at the school was assigned to her with the intention of providing higher quality instruction for students in the most need. The same decision was made at each of the other schools. Although, it was not always the case that the IA who showed the most promise at training had the higher levels of treatment quality, this did occur most of the time.

The same decision process was used to pair the GAs with student groups for ELLs. For every unit increase in treatment quality, ELLs gained 1.219 points fewer on LWID and .491 points fewer on WA. Two graduate students in special education served the 23 ELLs in a multiple baseline design. Though both instructors had relatively high levels of treatment quality, 15.3 and 18.4, for M.C. and L.K., respectively, the regular
classroom teacher for M.C.’s students held a doctorate in education with an emphasis on reading. It is possible that this teacher provided a higher quality of classroom instruction that L.K.’s students did not receive.

Overall, this model does not contribute any meaningful explanation for gain scores on LWID and WA for non-ELL and ELL students in this study. It does however, support the NRP’s (2000) meta-analysis on early reading intervention, which concluded that the best outcomes are found for treatment durations between 5 and 18 hours. Further, although the design used for ELLs precludes any definitive conclusion regarding the influence of treatment duration and quality, the results for treatment quality and gain scores for non-ELLs are supported by the Perepletchikova and Kazdin’s (2005) contention that although the relationship may be intuitive, the literature base is still mixed. Some studies suggest a positive relationship between treatment integrity levels and treatment outcomes (Gresham, 1989; Peterson, Homer, & Wonderlich, 1982), but several studies fail to support this hypothesis (Bein et al., 2000; Burke, 1996; Noell, Witt, Gilbertson, Ranier, & Freeland, 1997; Patton, 1988; Toffolo, 2000). The effects of varying levels of treatment integrity on outcomes is difficult to analyze since the quality with which IAs implement any treatments will naturally differ according to many factors such as opinions of the interventions and expertise with the protocol. In this study we could not experimentally manipulate the level of treatment quality, so we had to analyze it as a correlate of outcomes, rather than a cause. Also, since IAs and students were nested within schools, we were limited in the extent to which we could pair more or less skilled instructors with more or less skilled students. Ideally, we could randomize assignments, but again, in applied research in education this is rarely an option.
Research Question Three

What proportion of ELL and non-ELL treatment students will improve their DIBELS recommendations?

Most students improved their DIBELS recommendation status or remained the same from fall to spring. Half of all non-ELL treatment students improved, while 43% of non-ELLs improved. The remaining ELLs remained the same, while 42% of non-ELLs remained unchanged. It’s important to point out that three non-ELL students (7.8%) regressed. Overall, from fall to spring, non-ELLs improved their DIBELS recommendations. This is particularly impressive since the DIBELS benchmark assessments are not static evaluations of performance that produce grade- or age-level equivalents. Rather, the DIBELS were designed to be indicators of essential literacy skill acquisition with predictive validity for future performance in other key target areas. Recent research indicates that the importance of the alphabetic principle to later reading acquisition is much greater than expected. In a yet unpublished study Good (2006) evaluated the amount of variance on subsequent DIBELS indicators (specifically ORF at the end of first-grade) that could be explained by risk status, alphabetic principle skills (as measured by the NWF subtest), and progress trajectory. In two separate samples (n=156,777; n-1234), the middle of Kindergarten risk category accounted for 25% and 22% of the variance on ORF, respectively. Students’ alphabetic principle skills in the middle of Kindergarten contributed an additional 15% and 14% of the variance explained, and progress from the middle to end of Kindergarten contributed an additional 9% and 16% of the variance explained. By the end of Kindergarten, risk status and progress with alphabetic principle skills accounted for 49% and 56% of the variance explained.
explained in these two samples. This suggests that intervention in alphabetic principle skills in Kindergarten can have a sizable impact on later ORF scores. Therefore, interventions that alter the trajectory of progress and risk status may have beneficial implications on later reading development.

Although the same proportion of ELLs were in need of intensive intervention in both fall and spring, the proportion of ELLs with a strategic recommendation reduced and those at benchmark increased. These improvements are notable since due to the multiple baseline design used with these students they had a briefer intervention period. It is possible that more students would have moved from an intensive recommendation if they had longer durations of intervention.

These results are promising given that these students demonstrated considerable risk for reading failure at the beginning of the year. In contrast, their higher-performing peers, who did not exhibit risk in the beginning of the year, began to regress to strategic or even intensive recommendation levels. In fact, approximately 22% of the control group regressed to a recommendation of either strategic or intensive. For this 22% who were once at benchmark, the educational prognosis suggests that these students are now at risk of reading failure. Students in the treatment groups improved their recommendations, with many achieving benchmark. This suggests that participation in the intervention can contribute to remediating early reading deficits and reducing the achievement gap between students at risk and their peers.

Due to the ordinal nature of the recommendations, a ceiling or floor effect may impede the interpretation of these data. In other words, treatment students had more possibility for variability in their data because both intensive and strategic (and
benchmark for ELLs) recommendations were well represented in the fall. However, although some strategic students were represented among the control group, the majority was at benchmark. For the most part, the only direction for change possible was regression.

Research Question Four

As a measure of social validity, to what extent will instructors, school principals, and related personnel view the supplemental instruction as beneficial to their students?

Instructors. Often, unless general education classroom teachers find an intervention acceptable, it is unlikely to be continued, regardless of its effectiveness (Polloway & Bursuck, 1996). We have some anecdotal evidence to support this claim. One instructional assistant in this study had experience co-teaching the program with a doctoral candidate during the previous school year. At the conclusion of that study, the IA provided a very positive evaluation of the program and requested that the principal purchase a copy of the curriculum so that she may continue to work with at-risk students during the next year. She did continue to work with at-risk students this year as part of this study and it is noteworthy that she had one of the highest levels of treatment quality among all of the IAs.

Nearly all of the instructors who participated in this study provided a positive rating. Most agreed or strongly agreed that this intervention is an acceptable way for dealing with students’ reading difficulties and would be willing to use the procedure to improve these difficulties. Most also agreed or strongly agreed that the intervention was likely to be effective and result in permanent improvement. However, there was a spread of opinion about the extent to which the instructors agreed that they liked the procedures.
These results suggest that although the IAs recognize the intervention’s appropriateness for addressing reading problems and potential effectiveness, they were not as comfortable with the specific procedures used to address reading problems. A few lines were provided for IAs to provide comments, but few specific comments pertaining to procedures were given. One can only speculate about which aspects of the procedures the IAs found to be disagreeable; however, it is possible that the structured and scripted nature of the program may be implicated. Direct observations of their teaching revealed that many did not adhere to the script faithfully. A few IAs who had particularly low levels of adherence were frequently reminded to follow the script and provided a rationale for doing so. Still, even after a retraining session prompted primarily by low levels of adherence to protocol, some did not improve this aspect of their implementation. The NRP (2000) addressed this issue in their report by noting that even though scripted programs promote better outcomes overall, they are less likely to be accepted and programs that encourage unique teacher contribution and offer flexibility should be adopted.

In the current study there is reason to be cautious about encouraging instructors to deviate from the script. First, deviations from the protocol create obvious problems for the replication and dissemination of this study’s findings. It is not likely that one can replicate the procedures or the effect without some clear definition of what actually happened. Second, the potential effectiveness of the program is predicated on several iterations of field-testing and revisions based on the project Optimize data. The extent to which the program can “tolerate” deviations and retain its effectiveness is not known. Finally, with the exception of L.K. who had five years of teaching experience, the instructors in this study were not experienced teachers. Although principles of effective
instruction were explained and described in the training session, a one-hour introduction to these principles is not adequate to permit these instructors to make curricular decision in the absence of data to support those decisions.

One IA who deviated from the script rather frequently was S.W. It is interesting to note that her responses to the social validity questionnaire were also relative outliers. She disagreed that the intervention was an acceptable way for dealing with reading difficulties and that the procedures were acceptable. She also did not think that the intervention was likely to be effective. She strongly disagreed that she had an overall positive reaction to the intervention. During the final meeting to discuss the project at the end of the school year S.W. had a rather negative reaction to the project and indicated that she felt the students progressed better with a journaling activity rather than the ERI. She did not provide more specific information about her preference for journaling over a structured early reading intervention. Additional anecdotal information, however, indicated that S.W.’s negativity was not limited to the project but extended to other conditions within the school, as well.

School principals and related personnel. The three principals and two other school related personnel returned the completed social validity questionnaires. These responses were overwhelmingly positive. For the most part, all agreed or strongly agreed that the intervention was an acceptable way to deal with reading difficulties, was likely to be effective and result in a permanent improvement, and that the procedures were acceptable. Principal 3’s responses were not quite as positive as the others, but she generally was in agreement or neutral and did not communicate a negative opinion. Some offered comments about the intervention, and R.F.C. noted that consistency of
implementation was important. R.F.C. was the Reading First Coordinator at School 3, which was also the school at which S.W. worked. The consistency of implementation was the weakest at School 3 and the general attitude at the school was less enthusiastic than at the other schools. The direction of influence is not clear. That is, it may have been that the intervention itself was disliked and therefore, the instructors implemented it less consistently. Or it may have been that since there was no clear requirement that it was implemented daily, the students exhibited less progress and the instructors began to view the intervention as ineffective or weak. Comparatively, Principal 2 communicated clearly to her staff that the intervention was a priority and that daily implementation was required. Moreover, once it was clear that the number of children needing intervention exceeded available personnel resources, she altered the schedule to permit multiple groups of children to receive intervention. Principal 1 was not as strong in her commitment to consistent implementation, as she consented to only conducting intervention three times per week rather than daily, but still openly communicated that it was a priority for her IAs.

Research Question Five

To what extent will students rate the supplemental instruction as beneficial?

Overall the responses to the social validity questionnaire were not substantially different for ELLs and non-ELLs. The overwhelming majority of responses were positive and the students often mentioned that they enjoyed the games and earning prizes. Furthermore, they mentioned learning specific skills such as certain letter sounds and words, how to read books, and say letter names. Almost all students indicated that they liked being part of the group “a little” or “a lot”. They also enjoyed spending time with
their instructor and earning rewards or stars “a little” or “a lot”. Results for whether they liked learning new skills were mixed, with some students indicating that they didn’t like this aspect very much. Students also noted that the skills they learned helped them “a little” or “a lot” in school and at home, though students were more likely to report that the skills they learned helped them “a lot” at school, this was particularly true for the ELLs. A strong majority of the students agreed that they liked being in the group, learned important things, and use the skills they learned in other settings. However, when asked if they wished the group could last longer, the majority still responded “yes”, but this majority was less pronounced. However, it is possible that this question was confusing to students because some responses reflected that students interpreted the question differently. For example, some responded that they wished it could last forever, ten more minutes, two days, and so on.

Limitations of the Study

Training and Support

Time is very limited in urban schools. This is particularly true with recent mandates that require specified minimum hours of instructional time. However, the successful delivery of this intervention required considerable advance preparation and practice. The curriculum calls for high rates of active students responding and the instruction is most effective when delivered with a brisk pace. Becoming skilled with a scripted curriculum takes time and training for even the most experienced teacher. Some estimate that in order for a professional development model to be effective, repeated training and supports may be necessary for up to two years (Goldenberg & Gallimore, 1991).
Unfortunately one hour of training, which was the most time allotted by the school principals, is far from sufficient to guide and support new instructors. Given this seriously limited training time, we decided to provide instruction on the elements of effective instructional delivery that are not explicitly stated in the script. For example, the difference between attending to the task and actively responding was explained and demonstrated. Other demonstrations included immediate and specific error correction, praise, and pacing. We also felt it was important to model a lesson so the instructors could immediately practice in pairs, while we supported and guided their efforts. Obviously, much more time would have been necessary to build fluency with instructional delivery. In some cases, instructors improved the quality of instructional delivery after some practice, as evidenced by A.H.’s treatment quality graph (see Figure 4.7). However, most demonstrated medium to low levels of quality and those levels remained stable throughout the course of the study.

Time for weekly feedback and practice sessions was requested; however, other school related duties such as recess supervision precluded this possibility. The primary researcher attempted to offer feedback and suggestions as the IA was cleaning up after instruction and transitioning to the next activity. This was not the ideal context to support improvement of instructional delivery and may have been more effective if provided a structured time and place to conduct ongoing training sessions.

Consistency

Studies that report the number of weeks or days an intervention took place provide an indication of the span of time in which an intervention occurred. However, it may be more important to note the actual duration in minutes or hours in order to allow a
more specific analysis of the duration of treatment and the outcomes obtained. If duration information is reported in future research or can be extracted from existing research, it may be possible to evaluate the extent to which treatment duration actually influences the dependent measures. In this study, treatment duration accounted for very little of the outcome scores on word attack. For example, for non-ELL students, only 3% of the variance was explained and the relationship was negative although the coefficient was negligible (−.001). For letter word identification, more variance was explained (22%), but again, the coefficient was negligible (.008). Similarly mixed results were observed for the ELL students, although as previously discussed, those results are confounded by the lesser treatment durations required in a multiple baseline design.

It is possible that some factor other than total treatment duration contributes more to outcomes. In the Al Otaiba, Schatschneider, and Silverman (2005) study, the authors manipulated the consistency or frequency of the intervention rather than the total duration. The authors concluded that students who received intervention four days per week made statistically significant greater gains than the students who received treatment two days per week. It is not clear from this study if students made those gains due to the duration of treatment or the frequency. It may be interesting to replicate the study, but control for the total treatment duration. We could speculate that students, who practice a skill more frequently, rather than for a longer period, would make greater gains. A clear example can be found for musicians. Students learning to play an instrument are often taught to practice short passages everyday rather than fatiguing themselves by practicing for one long session only once per week. By practicing a skill less frequently, there is some time necessarily devoted to “catching up”. In the current study, some IAs did not
implement the intervention consistently. For example, S.W. would often skip intervention sessions for some groups for days at a time, and then work with students several days in a row to “catch them up”. However, since much time had passed between sessions, she often went back several lessons. This type of sporadic implementation is not evident from the duration data alone and to truly understand the effects of the intervention for these students, we may need to run the analysis on consistency, or even the number of lessons taught, rather than duration.

Another complication with the duration data was that since the instructors were not observed everyday, it is possible that the data they supplied was not accurate. The school principals attempted to convey support for the intervention and required its implementation to varying degrees. Further, when the primary researcher was not present to observe the instruction, those recorded intervention times cannot be verified. Although school principals verbalized support during group meetings and outlined intervention schedules, there is no evidence that they followed up by observing the intervention themselves. The instructors may have felt some anxiety about having to record their intervention times accurately if they were not implementing the intervention as often as their principals encouraged.

*Specification Error*

As stated in chapter 3, the selection of variables to include in the analysis is based on the theoretical knowledge base. Review of the current literature and the report of the NRP (2000) suggests that an explicit, intensive, and supportive curriculum that incorporates alphabetic principle skills such as phonemic awareness and phonics instruction is effective. Very little literature exists on more specific treatment delivery
elements that may affect outcomes, but there is a growing, albeit mixed, line of research that suggests the influence of treatment integrity. Even less literature exists that outlines the effects of varying durations of treatment, however, the NRP’s meta-analyses did find that effect sizes were the most robust for treatments lasting between 5 and 18 hours. This evidence supported the selection of these variables for this study, so the relatively weak models are not due to inclusion of irrelevant variables, rather it is more likely that other relevant variables were excluded, or the variables were not framed as accurately as they could have been. A related point is made above regarding examination of frequency data rather than duration.

When scientists refer to fields such as chemistry and physics as the “hard science”, some educational researchers may reply that there is no harder science than trying to conduct research in an applied setting with an infinite number of influential variables. The selection of the variables explored in this study is sound. However, clearly other variables explain some of the variance on the outcome scores. The quality and type of classroom instruction, exposure to reading related activities outside of the classroom, or receptive or expressive language levels may all have affected the outcome scores. In fact, there is literature to support the influence of various elements of language on reading ability (Roth, Speece, & Cooper, 2002). It is worth differentiating here between the types of goals pursued in educational research. While it is certainly a worthwhile endeavor to determine all the most relevant variables that contribute to reading outcomes, variance that is explained by current levels of functioning or scores on other tests is not information that ultimately influences how we teach children to read. It only tells us how well they need to score on tests. Intervention research, on the other hand, is concerned
with what educators can alter about their teaching that will affect how well students learn. After all, although many things do affect students’ performance, teachers need to be more concerned with the variables they can change. Therefore, although it may appear that little variance was explained on the whole, ultimately, the variance that was explained may have more practical application than other nonmanipulable variables.

Sample

Since randomization was neither preferable nor possible, the results of the study cannot be generalized to the greater population. With extreme populations such as special education, it is not ethical to randomly exclude children from a potentially effective treatment. So, in this study, children with substantial deficits were compared to a control group of students without such impairments. Although a more comparable control group would have been preferable, in some ways this comparison lends more credibility to the intervention. The outcomes obtained for the treatment students, particularly for the ELLs, were quite robust considering that they either approached or exceeded the outcomes for the control students. It would be expected that ELLs would not approach or exceed the scores of the non-ELL controls. However, the ELLs in this study not only made impressive gains, but may have effectively reduced the achievement gap.

Implications for Practice

*English Language Learners*

The results of this study clearly support the use of an English-language early reading intervention to improve early reading skills for ELLs. These findings appear to support Foorman and colleagues’ (2004) conclusions that since phonological processes are emphasized in early reading, rather than meaning-based skills, ELLs can benefit from
English-language early reading interventions. The ERI’s focus on phonemic awareness and alphabetic principle skills appears appropriate for teaching early reading skills to students who do not have English as their primary language. Further, the results for ELLs in this study support Sharon Vaughn’s (2006) claim that when students are provided explicit early reading intervention, these students can make equivalent, if not better, progress on measures of early reading skills.

Other authors (Gunn et al., 2005; Lesaux & Siegel, 2003) also point to the possibility of ELLs making better progress than their non-ELL peers when provided intervention that adheres to known effective instruction principles. As seen in figures 4.1 and 4.2 of this study, the ELLs did approach or exceed the scores of the non-ELL treatment students and the control students. It is particularly noteworthy that the ELLs slightly surpassed the control on the word attack subtest. The nature of this test, more so than the letter word identification subtest, requires students to use letter sound knowledge to sound out nonsense words. It is possible that ELLs were better able to respond on this test since it does not require previous knowledge of English vocabulary. That is, it is possible that non-ELL students may have an advantage on the LWID subtest since they could make fairly accurate guesses since these students have more extensive backgrounds with English vocabulary. ELLs may not have this same history and therefore are at a slight disadvantage. The use of nonsense words on the WA subtest does not present this same kind of bias.

The LWID subtest requires students to not only apply letter sound knowledge, as in the WA subtest, but extends that application to irregular English words. This may explain why ELLs fared the lowest of the three groups. However, the trajectory of
progress was still very impressive and considering that these students had the least
duration of intervention, it is possible that with more intervention they would have
matched or exceeded the scores of the other students.

*Instructional Assistants*

The constraint on resources, both personnel and financial, in urban schools is not
new. However, with the introduction of recent Reading First requirements, teachers’ time
is less flexible than ever. Teachers must adhere to specified curricula and specified
schedules for core reading instruction, including mandatory intervention periods. Yet, it
is not uncommon in the poorest urban schools that there may be more students who need
intervention than the teacher alone can accommodate. This is particularly true since the
diversity of students’ abilities and backgrounds, including a wide variety of primary
languages, make truly individualized intervention a prohibitive occurrence. The results of
this study are encouraging in the sense that ELLs can learn to read in English with an
English-language intervention, thereby saving costs associated with the hiring, training,
and retaining teachers or assistants with language backgrounds of all ethnicities
represented in the classroom. Still, the number of students who need intervention in early
reading skills is overwhelming.

This study demonstrates that one possible solution for this problem may be to use
IAs as interventionists. With careful and extensive training these personnel may be able
to provide much needed intervention services to students. In the previous study
conducted by Musti (2005), D.M. (School 1) served as a co-teacher with Musti to deliver
intervention to students at risk for reading failure. However, D.M. conducted only 11% of
the total intervention sessions independently. In the current study, D.M. conducted 100%
of the lessons independently and demonstrated a relatively high level of treatment quality (74.1%), which suggests that after some guided training and experience with a supplemental intervention IAs may be effective teachers for students at risk.

As discussed previously, it is important to ensure high quality training and ongoing support when using IAs in an instructional capacity. Furthermore, establishing support of the intervention program on part of the classroom teacher and school principal may be key to ensuring consistent delivery and maintenance of the intervention within the school. At times during this study, and during the previous study conducted by Musti (2005), priority for the intervention period was not respected, as IAs were often pulled away from intervention instruction for clerical duties or otherwise could not be freed on a consistent basis for instruction. However, these occurrences were much less frequent in the current study than in the previous one. As Musti points out, “In an ideal situation, the co-teaching format would have involved one or two weeks of training period, after which point the IA would have independently implemented the instruction. The students receiving instruction solely from the IA would have provided a more natural learning environment than to receive instruction from both the IA and the researcher. Furthermore, it is important to study if the ERI, implemented by regular school personnel, would produce similar findings” (p. 158). The findings of the current study produced similarly beneficial outcomes for students who received intervention from IAs alone, thereby supporting the premise that these school personnel can and should be used in a supplemental instructional capacity.
Directions for Future Research

Much remains to be done in early reading intervention research. However, the small emerging literature based on effective early reading instruction for ELLs, and the extent to which school personnel can be effective intervention leaders is promising. The ERI, which is approved by Reading First, is a commercially available effective early reading intervention that is easy to adapt to the regular classroom schedule. The current study extended research by Musti (2005), which demonstrated the intervention’s effectiveness for a small number of students, by expanding the findings to a greater number of students, including ELLs, and by turning over the complete responsibility of instruction to IAs.

Future research should pursue a longitudinal analysis of the findings by following up with the same cohort of 93 students through first and second grades. This type of analysis can provide valuable information regarding the extent to which the gains observed in this study can be maintained over time. Further, some guidelines could be established around the number of students who may continue to require secondary prevention supports, or even tertiary interventions.

The project PRIDE model (Bursuck, Smith, Munk, Damer, Mehlig, & Perry, 2004) may be a useful guide for future directions of research related to the outcomes of this study. In project PRIDE, students were screened for risk at three different schools in Kindergarten and provided some level of intervention based on those screening results. The students were followed during the following three years, as grades 1, 2, and 3 were added. The monthly or bi-monthly assessment data were the basis for decisions on which students should enter or exit the interventions. The current study would supply valuable
data to serve as the first year of this type of longitudinal study. The flexible entry and exit
data-based determinations should prevent school personnel from becoming overburdened
with the number of students entering the intervention. Also, as needed, some additions to
the intervention may be warranted. The ERI curriculum provides a sound basis for
alphabetic principle skills, and in later lessons, addresses oral reading fluency to some
degree. However, more emphasis on oral reading fluency may be beneficial for students
who do not naturally apply their alphabetic principle skills to connected text. In this case,
a brief, but intensive, oral repeated reading strategy may be appropriate. For ELLs, some
sight word peer-tutoring may help to improve their letter word identification scores by
increasing English-language vocabulary skills.

Finally, although the results in this study and literature base on treatment integrity
and outcomes are mixed, more research into the nature of this connection is needed.
Perepletchikova and Kazdin (2005) suggest that, “treatment integrity may be just a proxy
variable for other influences that may impact and even obscure the relationship between
integrity and outcome” (p. 367). In this study, the data collected reflected the extent to
which certain characteristics of treatment delivery were present. However, Perepletchikova
and Kazdin suggest that experience and motivation of the interventionist
may play a role, particularly if the intervention is time-consuming, difficult to deliver
faithfully, or the procedures are not acceptable to the interventionist. This, in part,
implicates levels of social validity as perhaps being related to outcomes as well. Further
investigations should collect more in depth information about the interventionists’
attitudes toward the intervention and the students. Based on recommendations from the
Perepletchikova and Kazdin paper, items may include: (a) the extent to which the
interventionist would be willing to spend his or her own time to plan for instruction, (b),
the extent to which the interventionist believes that students outcomes are strongly related
to instructional quality, (c) the amount of feedback and support the interventionist feels
was available, and (d) the extent to which the interventionist feels the training was
sufficient to prepare him or her to teach the curriculum.

Summary

This study examined the effectiveness of participation in ERI (Simmons &
Kame’enui, 2003) for 61 students at risk for reading failure. Students were identified as at
risk according to the DIBELS benchmark assessments and the word attack and letter
word identification subtests of the Woodcock Johnson Tests of Achievement. A total of
23 ELLs from two different schools and 38 non-ELLs from three different schools
participated in the study. Another objective of the study was to determine the extent to
which elements of treatment delivery, specifically treatment duration and quality,
accounted for variance in students’ outcome scores. Overall, participation in the ERI was
found to be beneficial not only for non-ELLs, but also especially for ELLs. Treatment
quality and duration did not account for a substantial proportion of the variance, but
warrants further investigation due to some design constraints of the study.

All students who received treatment made practically meaningful gains on critical
early reading skills as reflected by the mean gain scores on the WJ-III tests and the
intervention recommendations provided by the DIBELS benchmark assessments. In fact,
the ERI students approached or exceeded their more competent control peers on the word
attack subtest and although the treatment students did not exceed the LWID raw scores of
control students, their slopes of progress were clearly steeper, indicating that with time
and continued intervention, these students may surpass controls. Also, while some controls regressed from benchmark to strategic or even risk status throughout the course of the year, only three ERI students out of 61 regressed. None of the non-ELLs were benchmark in the beginning of the year, but 42.1% achieved benchmark by the end of the year. Similarly, the percentage of ELLs at benchmark increased by more than 30%.

This study provides evidence that an intensive, explicit, and supportive early reading intervention program, in this case, can substantially improve outcomes, and possibly longer term learning trajectories, of both non-ELLs and ELLs at risk for reading failure. Also, socially significant implications include the use of IAs to deliver this instruction and that ELLs can benefit from English-language reading interventions, thus eliminating the need to delay instruction until primary language literacy is established. These implications may alleviate some of the logistical and financial burdens as the outcomes-driven federal mandates increase demands on urban schools to meet or exceed ambitious achievement goals for all students. Finally, it is clear from this study that when effective interventions are implemented for students who present some of the greatest risks, such as achievement deficits and high poverty, their achievement can meet or exceed that of their non-risk peers, and the prognosis for future learning begins to brighten.
REFERENCES


Torgesen, J.K., & Bryant, B.R. (1994). *Phonological awareness training for reading.* Austin TX: PRO-ED.


APPENDIX A

PARENT LETTER FOR PARTICIPATION (ENGLISH)
January 11, 2006

Dear Parent:

I am a professor in the college of education at The Ohio State University. My graduate students and I will be conducting a research project in your child's school. We wish to see if the use of academic instruction in essential reading skills, particularly instruction in phonemic awareness and alphabetic principle, will bring about improvements in both academic achievement and school adjustment. The focus of our project is to improve the reading scores of students in Kindergarten and first grade. We will be using a prevention-based reading curriculum that provides instruction in some key areas of beginning reading.

We hope these strategies will prevent reading failure and help children to be more socially and academically successful in school. Your child’s classroom teacher and instructional aide will be closely involved in conducting the project, teaching and monitoring your child’s performance in reading skills. Your child will not be removed from any classroom instruction and will not lose any academic instructional time.

We also are requesting permission to videotape/photograph your child’s classroom. The purpose of these pictures is to demonstrate specific teaching and behavioral management strategies used by your child’s teacher. They will not be used to identify your child in any way. The tapes will be used in our teaching seminars to show other teachers how to implement these strategies. We need to demonstrate the use of these strategies in actual classrooms. If you do not consent to the photographs, we will locate your child outside the range of the camera but your child will remain in the classroom and continue with the classroom instructional activities.

We will review your child’s school records, and conduct periodic assessments to monitor his/ her performance throughout the study. Data collected on your child will include both academic and social behavior. If available, we will get from your child’s classroom teacher the district-wide test scores. All information collected about your child will be confidential. No one other than the researchers will use this information and your child will not be identified in any way to others.

Both at the middle and before the end of the school year, we may ask parents to complete a questionnaire on how effective you feel this project was on your child’s academic and social performance. We expect the questionnaire to take about 10 minutes to complete. We will also interview your child to determine how your child feels about the behavior
management procedure. This informal interview will take approximately 15 minutes to complete and it will not take away any of your child’s academic learning time.

We are requesting your permission so that we might use your child’s academic and social performance as data in this study. We also are asking permission to include your child in our classroom videotapes or photographs. Permission is purely voluntary and the decision not to permit this access will not affect the way your child will be treated or graded at school. Should you consent, please know that you can choose to withdraw your permission at any time during this project. If you have questions, please feel free to contact me at 292-7629. Thank you for your attention and cooperation.

Sincerely,

[SIGNATURE HERE]

Gwendolyn Cartledge, Ph.D.
Professor
APPENDIX B

PARENT CONSENT FOR PARTICIPATION IN EDUCATIONAL RESEARCH
(ENGLISH)
CONSENT FOR PARTICIPATION IN SOCIAL AND BEHAVIORAL RESEARCH

Protocol title: “Improving the School Success for Urban Learners.”

Protocol number: Pending

Principal Investigator: Gwendolyn Cartledge

I consent to my child’s participation in research being conducted by Dr. Gwendolyn Cartledge of The Ohio State University and her assistants and associates.

The investigators have explained the purpose of the study, the procedures that will be followed, and the amount of time it will take. I understand the possible benefits, if any, of my child’s participation.

I know that my child can choose not to participate without penalty to me and/or my child. If I agree to participate, I can withdraw my child from the study at any time, and there will be no penalty.

I consent to the use of videotapes and photographs. I understand that these pictures will only be used to demonstrate classroom teaching practices. My child will not be identified by name and my child will be depicted in these tapes in positive ways.

I consent to the use of the following information from my child’s school records and academic records: attendance, individualized education plan (if any), medical reports (if any), classroom test scores, and benchmark evaluations.

I have had a chance to ask questions and to obtain answers to my questions. I can contact the investigators at (614) 292-7629. If I have questions about my rights as a research participant, I can call the Office of Research Risks Protection at (614) 688-4792.

I have read this form. I sign it freely and voluntarily. A copy has been given to me.

Print the name of the participant: __________________________________________________

Date: _______________________________  Signed: ___________________________________

(Participant)

Signed: _______________________________

(Principal Investigator or his/her authorized representative)

Signed: _______________________________

(Person authorized to consent for participant, if required)

Witness: _______________________________

(When required)
APPENDIX C

PARENT LETTER FOR PARTICIPATION (SOMALIAN)
February 6, 2005

Waalidka qiimaha leh

Waxaan ahay professor kulliyada waxbarashada ee Ohio State University. Ardayda qalin jabineysa iyo anigaba waxaan sameeyneynaa mashruu baaris ah oo ilmahaaga dugsigiisa ah. Waxaan rabi laheyn in aan aragno hadii isticmaalka manhajka barashadiisa ay noqoto lagama maarmin fadlan akhriska, khaas ahaan barida ku dhawaaqida xarfaaha iyo muhiimada alphabetada, waxey keeni doontaa horumar dhinaca cilmiga ah iyo hagaajinta dugsiga. Xoog saaridda mashruuceeenaan waa hagaajinta dhibcaha akhrinta ardayga ku jira xanaanada iyo fasalka koowaad. Waxaan isticmaali doonaa kahortag ku saleesan manhajka akhrinta taasoo siinaysa baridda meelaha furaha u ah bilowga akhrinta. Waxaana rajeyeeynaa in ay xeadalahaas ka hortegi doonaan guuldaro akhrinta ah iyo ilmaha oo laga caawiyiyo dugsiga in ay aad ugu guuleystaan cilmiga iyo bulshadaba. Ilmaaga macallinkiiisa fasalka iyo gargaarihiisaba waxay aad u dhxegeli doonaan sameyntha mashruuca, barida iyo indha indheyntha ilmahaaga waxqabadkiisa xagga xirfadda akhriska. Ilmahaaga lagama saari doono fasalka waxbarashadiisa kamana dhumeysyo wakhtigii waxbarashada.


Waxaanu dib u fiirineeynaa diiwaanka ilmahaaga dugsigiis, waxana sameeyn doonaa qiimeyn xilliyeed lagu indha indheyn doono guud xayn ahaan ilmahaahaaga waxbarashadiisa. Haddii ay suurto gal tahay, waxaana ku heleeynao ilmahaaga macallinkiiisa ama gobolka meexa lagu keydiyo dhibcaha imtiixaanaadka. Warbixintaas la ururiiyey oo dhin oo ku saabsan ilmahaaga waxey noqoneysaa sir. Ma ogaan karto cid kale aan ka aheyn baarayaasha oo isticmaali doona, si kastaba ilmahaaga lama aqoonsanayo.

Labada xilli bartamaha iyo inta uusan sannad dugsiyeedka dhamaanin, waxaan waalidka weydiin doonaa in ay soo dheaemarka su’aalaha sida aay dareentyo ama uu wax tar u yahay meshruucaan fulintiisa xag bulsho ama xag tacliimeedba. Waxaan rajeyeeynaa in su’aalaha lagu dhameystiri doono ilaa 10 daqiqadood. Waxaan kale oo wareysa la

185
yeelan doonaa ilmahaaga in uu go’aansado sidii uu u dareemay habka maamulka dabeecada. Wareeyigaan aan rasmiga aheyn waxey qaadanaysaa in lagu dhameystiro qiyaastii 10 daqiicadood lagamana qaadi doono ilmahaaga waqtiga waxbarashadda.

Kani waa sannad waxbarasho badan leh waxaa suurtagal ah oo aan rabnaa in aan u sii wadno howlahaan sannadka dambe ilmahaaga, haddii ay sii socoto waxaan markale kaa cotsaneynaa ruqsad uu ilmahaaga kaga qeyb galo.


Daacadnimo ah

Gwendolyn Cartledge, Ph.D
Professor
APPENDIX D

PARENT CONSENT FOR PARTICIPATION IN EDUCATIONAL RESEARCH (SOMALIAN)
February 6, 2006

Ogolaanshaha ka qeybqaadashada bulshada iyo baarista dabeecadahooda.
Habka tirada: Laalan(Wali)
Baare maamule: Gwendolyn cartledge
Waxaan u ogolaaday ilmaheyga baaritaankii ey sameysay Dr. Gwendolyn cartledge oo ka socota Ohio State University kalkaaliyaasheeda iyo dadka la xiriira.
Baareyaasha waxey fasireyn ujeedada wax barashada, habkaan wuxuu sosomeedoono waqti.
Waxaan fahmay faa’iidooyinka laga heli karoo haduu ilmaheyga ka qeybqaato.
Waxaan ogolaaday in la isticmaalo videotapes iyo masawiro waxaan ogahay in masawiradaas loo isticmaaloon doono oo kali ah fasalka wax lagu barayo. Ilmaheyga lama aqoonsan doono magac ahaan.. Ilmaheyga wuxuu uga muuqan doona cajalada si hubaal ah.
Waxaan ogolaaday isticmaalka warbixintaan soo socota oo ka imaaneysa diiwaanka dugsiga ilmaheyga iyo diiwaanka manhajka, xaadirinta, qorsan tacliin shaqsiyeed(Hadii ey jirto,) Warbixin caafimad(Hadii ey jirto) dhibcaha intixaanka fasalka dhexdiisa ah, iyo qiyaasta qiimeyntooda.
Taariiqda siiixiay: ____________________________
Ka qeybqaate
Maamulaha (baaritaanka asaga/ayada wakiilka amarbixiyaha.) siiixiay:
(Amarbixiyihii ogolaaday ka qeybqaadashada, hadii loo baahday)

Marqaati: ____________________________
APPENDIX E

PARENT LETTER FOR PARTICIPATION (SPANISH)
Enero 31 del 2006

Estimado padre/madre:

Yo soy una profesora en el colegio de educación en la universidad “The Ohio State University.” Mis estudiantes de escuela graduada y yo estaremos conduciendo un proyecto de investigación en la escuela de su hijo(a). Desearíamos ver si el uso de instrucción académica en destrezas esenciales de lectura, particularmente instrucción en conocimiento de la fonética y principios alfabéticos, traerá mejorías en ambos logros académicos y ajustes escolares. El foco de nuestro proyecto es el mejorar los puntajes de estudiantes en el jardín de niños y primer grado. Estaremos usando un plan de estudios de lectura basado en prevención que provee instrucción en algunas áreas clave del principio de lectura.

Esperamos que estas estrategias prevendrán reprobio en lectura y ayudarán a los niños a ser más triunfantes académica y socialmente en la escuela. El maestro(a) del salón de su hijo(a) y asistente de instrucción estarán envueltos muy de cerca en la conducción del proyecto, enseñando y monitorizando el funcionamiento de su hijo(a) en destrezas de lectura. Su hijo(a) no será removido(a) de ninguna instrucción en el salón de clases y no perderá ningún tiempo académico.

También estamos pidiendo permiso para fotografiar/grabar en video el salón de clase de su hijo(a). El propósito de estas fotos es el demostrar estrategias de manejar conducta y enseñanza usadas por el/la maestro(a) de su hijo(a). No serán usadas para identificar a su hijo(a) en ninguna manera. Las cintas serán usadas en nuestros seminarios para enseñar otros(as) maestros(as) como implementar estas estrategias. Necesitamos demostrar el uso de estas estrategias en salones de clases en la actualidad. Si usted no consiente a las fotografías, colocaremos a su hijo(a) fuera del alcance de la cámara, pero su hijo(a) permanecerá en el salón de clases y continuará con las actividades educativas del salón de clases.

Revisaremos los expedientes escolares de su hijo(a), y conduciremos pruebas periódicas para monitorizar su funcionamiento por el curso del estudio. Si están disponibles, obtendremos los puntajes de pruebas del distrito por medio del maestro(a) de su hijo(a). Toda la información coleccionada acerca de su hijo(a) será confidencial. Nadie aparte de los investigadores usarán esta información y su hijo(a) no será identificado(a) en ninguna manera a otros.

A mediados y final del año escolar pediremos que padres completen un cuestionario acerca de cuán efectivo usted siente que este proyecto fue en el funcionamiento social y
académico de su hijo(a). Puede esperar que el cuestionario tome como 10 minutos para completar. También entrevistaremos su hijo(a) para determinar cómo su hijo(a) se siente acerca del procedimiento de manejo de conducta. Esta entrevista informal tomará aproximadamente 10 minutos y no quitará ningún tiempo académico de su hijo(a). Este es un estudio de múltiples años así que será posible que deseemos continuar estas actividades el próximo año con su hijo(a). Si es así, nuevamente pediremos su permiso para la participación de su hijo(a).

Estamos pidiendo su permiso para usar los datos de funcionamiento académico de su hijo(a). También estamos pidiendo permiso para incluir su hijo(a) en cintas de video o fotografías. El permiso es puramente voluntario y la decisión de no permitir este acceso no afectará la manera en que su hijo(a) es tratado(a) o calificado(a) en la escuela. si usted consiente, por favor sepa que usted puede escoger declinar su permiso en cualquier momento durante este proyecto. Si usted tiene preguntas, por favor síntase libre de contactarme en el (614) 292-7629. Gracias por su atención y cooperación.

Sinceramente,

Gwendolyn Cartledge, Ph.D.
Profesora
APPENDIX F

PARENT CONSENT FOR PARTICIPATION IN EDUCATIONAL RESEARCH
(SPANISH)
CONSENTIMIENTO PARA LA PARTICIPACIÓN EN INVESTIGACIÓN SOCIAL Y DEL COMPORTAMIENTO

Título del protocolo: “Mejorando el Triunfo Escolar para el Estudiante Urbano”

Número del protocolo: Pendiente

Investigador(a) Principal: Gwendolyn Cartledge

Consiento a la participación de mi hijo(a) en la investigación siendo conducida por Dra. Gwendolyn Cartledge del colegio The Ohio State University junto a sus asistentes y socios(as).

Los investigadores me han explicado el propósito del estudio, los procedimientos que han de ser seguidos, y el tiempo que ha de tomar. Entiendo los beneficios posibles de la participación de mi hijo(a), si alguno.

Sé que mi hijo(a) puede escoger no participar sin penalidad alguna a mi o mi hijo(a). Si estoy de acuerdo a participar, puedo remover mi hijo(a) en cualquier momento, y no habrá penalidad.

Consiento al uso de la siguiente información del expediente de mi hijo(a) y expediente académico: asistencia, plan educativo individualizado (si alguno), reportes médicos (si alguno), puntaje de pruebas del salón de clases, y evaluaciones estandarizadas.

He tenido la oportunidad de hacer preguntas y obtener respuestas a mis preguntas. Puedo contactar los investigadores en el (614) 292-7629. Si tengo preguntas en cuanto a mis derechos como participante de la investigación, puedo llamar la Oficina de Protección de Riesgos de la Investigación (Office of Research Risks Protection) en el (614) 688-4792.

He leído esta forma. La he firmado libre y voluntariamente. Una copia se me ha sido entregada.

Imprima el nombre del participante: _________________________________________________________

Fecha: ______________________________  Firmado por: ______________________________

(Firmado por: _________________________ Firmado por: ________________________________

(Investigador(a) Principal o su representante autorizado(a))

(Persona autorizada a dar consentimiento por el/la participante, si requerido)

Testigo: _____________________________

(Cuando requerido)
Soy de la universidad “The Ohio State University.” Estoy aquí para trabajar contigo y algunos de los otros estudiantes en tu clase en maneras de ayudarte a ser un muy buen lector(a). Haremos estas actividades por (diga los minutos exactos) por (diga el número exacto de veces) veces cada semana.

¿Te gustaría participar en estas actividades? (respuesta del estudiante)

Espero que te guste hacer esto, pero si en algún momento no quieres participar en estas actividades, puedes pedir que terminemos. Si escoges no continuar, estará bien y nadie se molestará contigo. No afectará tus notas/calificaciones en ninguna manera.

¿Todavía quieres continuar? (respuesta del estudiante)
APPENDIX G

PRINCIPAL LETTER OF AGREEMENT
October 12, 2005

Gwendolyn Cartledge, Ph.D.
The Ohio State University
356 Arps Hall
1945 High Street
Columbus, Ohio 43210

Dear Dr. Cartledge:

I have had a chance to review the abstract and to discuss the proposal for the P-12 project: Reducing Special Education Risk for Young At-Risk Urban Learners. I understand that this project is designed to reduce reading failure by providing small group intensive instruction for students who show risk with initial assessments. Specifically, you propose to train and provide ongoing assistance to instructional assistants who will provide intensive instruction to these students in small groups. The phonemic and phonological awareness instruction will follow the Early Reading Intervention curriculum. My staff and I recognize the need for special interventions for these students. Furthermore, we believe that progress for this target group would bring about corresponding gains for all of our students.

In addition to the training provided by the instructional assistants, I understand that there will be collaboration with the classroom teachers in the following ways. Teachers will be instrumental in selecting students, in receiving information from your staff on the nature of this instruction, on scheduling students for interventions, and in consulting with you and your staff on student progress.

I feel this project offers special advantages in that it provides the opportunity to give many of our struggling students that special foundation they need to become competent readers and successful students. It also provides special training to our instructional assistant that will maximize their talents in the interest of our students. Finally, if the effects of this intervention are as desired, we will be able to continue this intervention in subsequent years with our staff.

I am very supportive of this project and am pleased that it has been funded. I feel this expanded project holds much promise and look forward to its implementation at our school.

Sincerely,

[Signature Here]

Principal
APPENDIX H

SAMPLE ERI LESSON
Lesson 15

New Sound /t/
Review Sounds /m/, /p/, /f/, /k/
Key Phonologic Skill Isolate Initial Sound

Phonological Awareness and Alphabetic Understanding

<table>
<thead>
<tr>
<th>Activity</th>
<th>Materials</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Alphabetic Introduce /t/</td>
<td>Alphabetic Card: /t/</td>
<td>1-2 mins</td>
</tr>
<tr>
<td></td>
<td>Letter Card: /t/ (one per child)</td>
<td></td>
</tr>
<tr>
<td>1. Integrated Phonologic/Alphabetic</td>
<td>Picture Choices: 15-1, 15-2, 15-3 Teacher Resource Package 1</td>
<td>4-5 mins</td>
</tr>
<tr>
<td>Which picture begins with /t/?</td>
<td>Alphabet Card: /t/</td>
<td></td>
</tr>
<tr>
<td>2. Integrated Phonologic/Alphabetic</td>
<td>Sound Match Cards: 15a – 15e (one card per child) Teacher Resource Package 1</td>
<td>4-5 mins</td>
</tr>
<tr>
<td>Sound Match with /k/, /f/, /p/, and /m/</td>
<td>Game markers (four per child/ not provided)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alphabet Cards: /c, /p, /m, /f/</td>
<td></td>
</tr>
<tr>
<td>3. Alphabetic Reintroduce /t/; Be the Teacher game</td>
<td>Alphabet Cards: /t, /m, /p, /f, /c/</td>
<td>4-5 mins</td>
</tr>
</tbody>
</table>

Writing and Spelling

| 5. Writer’s Warm-up Letter Mission: Review /t, /c, /f, /p, /m/ | Letter Mission (one per child) Student Activity Book 1, p. 48 or 49 Tracing Cards: /t, /c, /p, /m, /f/ | 2-3 mins |
| 6. Integrated Phonologic/Alphabetic Match a Picture’s Initial Sound; Introduce /t/ | Picture Cards: all /t/ cards Letter Cards: /t, /f, /m, /c, /p/ | 2-3 mins |
| 7. Phonologic/Spelling Letter Writing Game: Review /t, /c, /f, /p, /m/ | Letter Writing Game (one per child) Student Activity Book 1, p. 50 or 51 Game markers (one per child/ not provided) Game Board 3 Number cube (not provided) | 6-8 mins |

Lesson 15 Overview

Note. From Scott Foresman Early Reading Intervention, Simmons & Kame’enui, 2003.
### Activity 1: Introduce Letter Name and Sound

**Introduce t /t/**

Objective: Children learn and trace t /t/  

Time: 1-2 minutes

<table>
<thead>
<tr>
<th>To Do</th>
<th>To Say</th>
</tr>
</thead>
</table>
| **Introduce letter name** | Hold up the t Alphabet Card | Model  
|  |  | Lead  
|  |  | Test  
|  |  | The name of this letter is t.  
|  |  | Say the name with me.  
|  |  | What is the name of this letter?  |

| **Introduce letter sound** | Continue holding up the t Alphabet Card | Model  
|  |  | Lead  
|  |  | Test  
|  |  | The sound for this letter is /t/, the tip of your tongue touches above your top teeth. Say /t/ and feel the tip of your tongue touch above your top teeth.  
|  |  | Say the sound for this letter?  
|  |  | What is the sound for this letter?  |

| **Test knowledge of letter name and sound** | Continue holding up the t Alphabet Card | Test  
|  |  |  
|  |  | What is the name of this letter?  
|  |  | What is the sound for this letter?  |

**Ongoing Assessment**

*If..... children make an error, then.....tell them the letter name or sound and have them repeat it. Repeat the test.*

| **Model tracing t** | Distribute the t Letter Cards. Hold up the t Alphabet Card. | Model  
|  |  | Lead  
|  |  | Watch. I’ll trace the letter t.  
|  |  | Now you trace the letter t three times with your finger. Say /t/ each time you trace the letter.  |

**Ongoing Assessment**

*If..... children make an error, then.....put your hand over their hand and guide them to trace the letter. Then have children try to trace the letter on their own. Repeat as necessary.*
Activity 2 Isolate Initial Sound

Which Picture Begins with /t/?
Objective: Children isolate initial /t/ and connect sound to letter

To Do | To Say
---|---
Model names of pictures | Gather the picture choices. Place Picture Choice 15-1 on the table. Point to fork.
Model | This is fork. What is this? Continue with the remaining pictures: map, teeth, cup, toe, fin. Then test children on all the picture names. For each picture, ask: What is this?

Introduce the game Which Picture Begins with /t/?

Now we are going to play a game. I’ll show you a letter. You’ll find the picture that begins with the sound for that letter.

Model the game Hold up the t Alphabet Card

Model | The name of this letter is t. The sound for this letter is /t/. Remember, when you say /t/, the tip of your tongue touches above your top teeth. Say /t/ and feel the tip of your tongue touch above your top teeth. What is the sound for this letter? My turn. I’ll show you how to play. Point to each picture in the top row and say: This is fork, map, teeth. I’ll find the picture that has the first sound /t/. Point to the t Alphabet Card. Teeth begins with /t/. Exaggerate the /t/ in teeth. Exaggerate the first sound and say the word: /t/, teeth. Model one example with the bottom row: cup, toe, fin.

Test Model

(Activity 2 continued)
| Play the game to test knowledge of /t/ | Hole up the *t* Alphabet Card. Display Picture Choice 15-2 (*tire, moon, can; maze, fish, tail*). Cover up the bottom row. Give individual turns | Test What is the **sound** for this letter? |

**Ongoing Assessment**

*If..... children make incorrect responses, then....model the correct answer. Review the sound production cue. Have children repeat the correct answer. Go back to the example a second time. If this activity is very difficult, use two pictures.*
Activity 3 Isolate Initial Sound

Sound Match with /k/, /f/, /p/, and /m/

Objective: Children isolate picture’s initial sounds and to letter

Time: 4-5 minutes

<table>
<thead>
<tr>
<th>To Do</th>
<th>To Say</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduce the game Sound Match</td>
<td>Gather Sound Match Cards 15a – 15e and the game markers.</td>
</tr>
<tr>
<td>Test children on how to play the game</td>
<td>Distribute the Sound Match cards and game markers to children.</td>
</tr>
<tr>
<td></td>
<td>Hold up the c Alphabet Card and point to it.</td>
</tr>
<tr>
<td>Play the game to test knowledge of initial sounds</td>
<td>Repeat the game procedure for /m/, /p/, and /f/.</td>
</tr>
</tbody>
</table>

Today we’re going to play Sound Match. I’ll show you a letter. Then you’ll find the picture on your Sound Match card that begins with the sound for that letter and place a marker on it. Watch carefully because I’m going to show you four different letters. Each person has a different Sound Match card, so we won’t now who’s going to win.

Have children review the picture names: Before we start, whisper to yourself the name of each picture to be sure you know each one. Raise your hand if you don’t know one and I’ll tell it to you. The picture names are listed below:

Test 

What is the sound for this letter? Find the picture on your Sound Match card that begins with /k/ and put a marker on it. Have each child name the picture he or she put the marker on. Confirm correct responses and have the whole group repeat the word: Yes, coat begins with /k/. Everybody, say /k/, coat.

Let’s continue playing. If you fill up a whole row or column, call out “Sound Match!” When a child wins, have everyone clear off their Sound Match Cards. Collect the cards, mix them up, and redistribute them. Play a second round, testing the sounds in the following order: /f/, /k/, /m/, /p/.

Sound Match Picture Names

15a: tub, tent, coat/ tape, tie, mitt/ paw, toe, fish
15b: tie, tag, pin/ tow, tub, tail. Fork, cap. Mouse
15c: fire, /op, pan/ tag, tape, tent/ tie, cat, toe
15d: cut, tae, tail/ man, tag, tub/ fan, tie, pie
15e: moon, cup, tub/ tent, foot, tag/ tail, pig, toe
## Activity 4  Review Letter Name and Sound

### Reintroduce t /t/; Be the Teacher Game

**Objective:** Children practice letter name and sound: t /t/

**Time:** 4-5 minutes

<table>
<thead>
<tr>
<th>To Do</th>
<th>To Say</th>
</tr>
</thead>
</table>
| **Introduce the letter name** | Model | The name of this letter is **t**.  
|                        | Lead  | Say the name of this letter with me: **t**.  
|                        | Test  | What is the name of this letter?  |
| Hold up the **t** Alphabet Card |                                  |
| **Introduce letter sound** | Model | The sound for this letter is /t/, the tip of your tongue touches above your top teeth. Say /t/ and feel the tip of your tongue touch above your top teeth.  
|                        | Lead  | Say the sound for this letter?  
|                        | Test  | What is the sound for this letter?  |
| Continue holding up the **t** Alphabet Card |                                  |
| **Test knowledge of letter name and sound** | Test | What is the name of this letter? What is the sound for this letter?  |
| Continue holding up the **t** Alphabet Card |                                  |

**Ongoing Assessment**

If….. children makes an error, then….tell them the letter name or sound, have them repeat the name or sound, and return to the letter a second time.

Ask individuals: What is the name of this letter? What is the sound for this letter?

**Introduce the game Be the Teacher**

Choose a child to be the “teacher.” Give him or her the **m, p, f, c, and t** Alphabet Cards.

Tell the “teacher” to hold up the first Alphabet Card. Have the “teacher” ask: What is the name of this letter? What is the sound for this letter? Allow the group to respond to each question. If the group answers correctly, have the “teacher” place the Alphabet Card on the table.

**Ongoing Assessment**

If….. the group makes an error, then….prompt the “teacher” to tell the letter name or sound and place the Alphabet Card at the bottom of the pile.

Continue until all the letters and sounds have been identified correctly. Then have the “teacher” give individual turns.
# Activity 5 Writer’s Warm-Up

**Letter Mission: Review t, c, f, p, m**
Objective: Children practice writing letters.

**Time:** 2-3 minutes

<table>
<thead>
<tr>
<th>To Do</th>
<th>To Say</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduce the activity</strong></td>
<td>Distribute a Letter Mission to each child.</td>
</tr>
<tr>
<td><strong>We’re going on a letter mission. I’m going to tell you the name of a letter and you’re going to write it. Let’s see if we can help the dog get the bone!</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Model the activity</strong></td>
<td>Hold up a Letter Mission.</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td>Let’s do one together. The first letter is ( t ). Watch as I start at the dot and write the letter ( t ). Write a ( t ) on your sheet.</td>
</tr>
<tr>
<td><strong>Lead</strong></td>
<td>Now you write the letter ( t ). Start at the dot and write the letter.</td>
</tr>
<tr>
<td><strong>Lead the activity to test writing ( t, c, f, p, ) and ( m )</strong></td>
<td>Continue naming letters for children to write</td>
</tr>
<tr>
<td><strong>Test</strong></td>
<td>Now write the letter ( c ). Reinforce the group on the letter name: What is the name of the letter you just wrote? Continue with the following: ( p, m, t, p, f, c )</td>
</tr>
</tbody>
</table>

**Ongoing Assessment**
If..... the children write the wrong letter or don’t remember a letter, then....show them the tracing card of the correct letter and model tracing the letter. Have them write the letter.
Activity 6 Connect Sound to Letter

Match a Picture's Initial Sound; Introduce t
Objective: Children connect initial sound to letter.

| Time: 2-3 minutes |

<table>
<thead>
<tr>
<th>To Do</th>
<th>To Say</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduce the activity</strong></td>
<td>Gather the picture cards and the letter cards.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Model the activity</strong></th>
<th>Place the tire Picture Card on the table.</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>This is tire. The first sound in tire is /t/. Place the t and f Letter Cards under the picture.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I'm going to choose the letter that matched this picture's first sound. Point to the correct letter and say: t is the letter for the sound /t/ like the /t/ in tire.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Test knowledge of initial sound /t/</strong></th>
<th>Display the top Picture Card.</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>This is top. What is this? The first sound in top is /t/. What is the first sound in top? That's right. /t/ is the first sound in top. Place the t and m Letter Cards on the table. Your turn. Choose the letter that matches the picture's first sound. Think. Which letter stands for the /t/ in top? Do you think you know?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Call on a child to point to the correct letter: What's the name of the letter you pointed to? What is the sound for that letter? That's right; t is the letter for /t/ like the /t/ in top. Reinforce the letter name and sound with the group: Everyone, what's the name of the letter? What's the sound for the letter? Continue with the remaining cards.</td>
</tr>
</tbody>
</table>

**Ongoing Assessment**

If.....children make an error, then.....model the answer, have them repeat it, and return to the picture card a second time.
Activity 7 Connect Sound to Letter

Letter Writing Game: Review t, c, f, p, m
Objective: Children connect initial sound to letter

**Time:** 6-8 minutes

<table>
<thead>
<tr>
<th>To Do</th>
<th>To Say</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduce the letter name</strong></td>
<td>Hold up a Letter Writing Game.</td>
</tr>
<tr>
<td><strong>Introduce the Letter Writing Game</strong></td>
<td>Distribute a Letter Writing Game to each child</td>
</tr>
<tr>
<td><strong>Model the game</strong></td>
<td>Place a Letter Writing Game to each child</td>
</tr>
<tr>
<td><strong>Play the game to test children on t, c, f, p, and m</strong></td>
<td>Dictate words from the Word Bank below. Vary the order.</td>
</tr>
</tbody>
</table>

After everyone has written the letter, have one child roll the number cube and advance his or her marker on the game board. Continue the game until all the letters have been written.

**Word Bank**

tie cap feather paw moon
If.....children make an error, then.....model the letter name and sound, have children repeat them, and have children write the correct letter.

Adapted by S. Musti-Rao (2005)
APPENDIX I

TREATMENT QUALITY CHECKLIST
**Procedural Integrity Checklist**

Implementer: _______________________ School: _______________________ Observer: _______________________

Date: ____________ Time: ____________ Lesson: ______ Student names: _______________________

---

**Instructions:**

1. Indicate the extent to which the implementer performs each of the following steps by checking the appropriate box.
2. Write comments, observations, or suggestions for improvement in the space provided.

* Red column indicates that the step was not implemented, green columns indicate that the step was implemented.

<table>
<thead>
<tr>
<th>Follows script</th>
<th>Never (0)</th>
<th>Rarely (1)</th>
<th>Mostly (2)</th>
<th>Always (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Models the skill, provides practice before individually assessing student on skill (For example, says “The name of this letter is “m.” Say the name of the letter with me: “m.” What is the name of the letter?”)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>Uses manipulatives (e.g., alphabet cards, picture cards) as outlined in the lesson</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>Provides immediate feedback/ error-correction (Example: says “Good” for correct responses, and provides correct response for incorrect responses)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>Uses appropriate pacing (e.g., fast for review skills, provides thinking pause on more difficult skills)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>Delivers error-correction according to script</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>Provides reinforcement (star card) approximately once every 5 minutes for appropriate behavior</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
</tbody>
</table>

**Session Integrity = Number of steps completed X 100**

**Quality of implementation (scale: 0-21)**

Notes/ Comments: _______________________

---

209
APPENDIX J

DAILY TIME LOG
ERI DAILY TIME LOG

Date: ______________ Group:______________ Lesson: ____
Exact time (min:sec) ____________ Reteach? ________________

Date: ______________ Group:______________ Lesson: ____
Exact time (min:sec) ____________ Reteach? ________________

Date: ______________ Group:______________ Lesson: ____
Exact time (min:sec) ____________ Reteach? ________________

Date: ______________ Group:______________ Lesson: ____
Exact time (min:sec) ____________ Reteach? ________________

Date: ______________ Group:______________ Lesson: ____
Exact time (min:sec) ____________ Reteach? ________________

211
APPENDIX K

SOCIAL VALIDITY QUESTIONNAIRE (PRINCIPALS and INSTRUCTORS)
Date: ______________

Intervention: Early Reading Instruction (ERI)

Instructional Assistant: ________________  Grade Level: K

Please complete the items listed below by circling only one answer that best indicates how you feel about the intervention. Please read the items carefully because a circle accidentally placed on one number rather than another may not represent the meaning you intended.

<table>
<thead>
<tr>
<th>Level of Agreement</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

To what extent do you agree or disagree with the following statements? (Circle only one number on each line)

1. I find this intervention to be an acceptable way of dealing with students’ reading difficulties.  

   | 1 | 2 | 3 | 4 | 5 |

2. I would be willing to use this procedure if I had to improve a student’s reading difficulties.  

   | 1 | 2 | 3 | 4 | 5 |

3. I believe that it would be acceptable to use this intervention without students’ consent.  

   | 1 | 2 | 3 | 4 | 5 |

4. I like the procedures used in this intervention.  

   | 1 | 2 | 3 | 4 | 5 |

5. I believe this intervention is likely to be effective.  

   | 1 | 2 | 3 | 4 | 5 |

6. I believe the student will experience discomfort during the intervention.  

   | 1 | 2 | 3 | 4 | 5 |

7. I believe this intervention is likely to result in permanent improvement.  

   | 1 | 2 | 3 | 4 | 5 |

8. I believe it would be acceptable to use this intervention with students who can not choose interventions for themselves.  

   | 1 | 2 | 3 | 4 | 5 |

9. Overall, I have a positive reaction to this intervention.  

   | 1 | 2 | 3 | 4 | 5 |
If there is anything else you would like to tell us, please do so in the space provided below:

………………………………………………………………………………………………
………………………………………………………………………………………………
……………………………………………………………………………………………...
APPENDIX L

SOCIAL VALIDITY QUESTIONNAIRE (STUDENTS)
Student Participant Interview

Student Participant: _______________       Interventionist: _______________________

Interviewer: _____________________        Date: _______________________________

1. Did you like being in this special group/ program? _____ Yes _____ No

2. What did you like best about this special group/ program? _______________________
________________________________________________________________________

3. Did you like…………                  Not Much  A Little  A Lot
   Being part of the group?            ☹       ☺       ☺☺☺
   Spending time with the Ms./Mr. _____ in
       the group?                     ☹       ☺       ☺☺☺
   Earning rewards/ stars?            ☹       ☺       ☺☺☺
   Learning new skills?              ☹       ☺       ☺☺☺

4. Do you feel you learned important things? _______ Yes ______ No
   If yes, What is/ are the most important thing(s) you learned?
________________________________________________________________________

5. Did you learn things that will ………
   Help you do better work in school? ☹       ☺       ☺☺☺
   Help you at home?                 ☹       ☺       ☺☺☺

6. Do you use the skills that you learned in the special program/ group? ____Yes ___ No
   If yes, where do you use these skills? ________________________________

7. Do you wish our group/ program could have lasted longer? (Clarify…..meet for more
   time, like for another week?) _______ Yes ______ No
   If yes, how much longer would you like to have met? ____________________________

8. Is there anything else you would like to tell me about your special group/ program?
________________________________________________________________________

Created by S. Musti-Rao (2005)