The Effects of a Computer-Assisted Reading Program on the Oral Reading Fluency, Comprehension, and Generalization of At-Risk, Urban Second-Grade 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

Starr E. Keyes, M.Ed.

College of Education and Human Ecology

The Ohio State University

2010

Dissertation Committee:

Dr. Gwendolyn Cartledge, Advisor

Dr. Ralph Gardner, III

Dr. Sheila Alber-Morgan
This study examined the effectiveness of a computer-assisted reading program, Read Naturally (RN), on the oral reading fluency (ORF), comprehension, and generalization of second graders who were at risk for reading failure. Six students received the intervention (RN) three to four times a week for approximately 7 to 12 weeks. All six students were trained to use the RN program, and were able to advance through the sequence of instructional activities with little to no assistance. The instructional sequence included: key words, cold timing, read along, practice reading, a comprehension quiz, pass timing, and word retell.

A multiple-baseline across participants design with embedded changing criterion tactics was used in this study. This design allowed the experimenter to note the effects of increasing the goal criterion (in correct words per minute [CWPM]) on the participants’ ORF. The experimenter increased each participant’s goal (CWPM) based on individual performance. Tiers one and two had six criterion changes over the course of the study, and tier three had five criterion changes. The dependent variables included: CWPM and incorrect words per minute (IWPM) on RN passages and two measures of generalization (AIMSweb and Direct Instruction [DI]); as well as the number of words re-told in a minute and the number of comprehension questions answered correctly on all three types
of stories (i.e., RN, AIMSweb, and DI). Pre- and post-test measures were given to
determine the effects of the RN intervention on each participant’s overall reading ability.

Results revealed that the RN program was effective at increasing the ORF of all
six participants on the RN treatment stories and AIMSweb generalization stories. Five of
the six participants increased their ORF on the DI generalization stories over baseline.
The two measures of comprehension revealed mixed results. Only half of the participants
showed increased performance on the percentage of comprehension questions answered
correctly, but all six participants increased the number of words re-told per minute from
baseline to intervention on all three types of stories. These results support the use of
computer-assisted reading instruction and repeated reading activities. This study extends
current research as each participant’s ORF treatment goal was increased based on
individual performance and shown to have a positive effect on their ORF on the
generalization measures. Implications, limitations, and future research are discussed.
Dedicated to my mom…

You are an amazing woman who has taught me more about life than you will ever know.

Words cannot express my gratitude for everything you have done for me.
ACKNOWLEDGEMENTS

First, I would like to thank God, for without Him, none of this would be possible. I would also like to thank Dr. Cartledge, my advisor, for being the best role-model anyone could ever ask for in academe and life. I would not have been able to complete this program without her guidance, love, and support. Thank you to Dr. Gardner and Dr. Morgan for serving on my candidacy and dissertation committee. You are great professors and people. Also, thank you to the faculty in the Special Education program. I have learned a great deal from you which I look forward to teaching to others.

Next, I would like to thank the administrators, staff, and students at the elementary school where this study was conducted. Thanks for allowing me to work in your school over the last two years, it was a pleasure. Jeff Banks, thank you for providing me with another laptop; you saved me! Porsha Robinson-Ervin, thank you for your support throughout the course of the study and the school year.

Last, but not least, I would like to thank my family and friends for your encouragement these last three years. You were there for everything, and I love you for that. Most importantly, my mom, my friend…Thank you for raising me so well and for providing me with anything I ever needed. We finally made it!
VITA

June 13, 1980.............................................. Born- Toledo, OH

June 14, 2002............................................. B.A. in Psychology,
The Ohio State University
Columbus, OH

May 8, 2005.............................................. M.Ed. Special Education
The University of Toledo
Toledo, OH

2002-2007.............................................. Teacher
Lake Erie Academy
Toledo, OH

2007-current ....................................... Graduate Trainee
The Ohio State University
Columbus, OH

PUBLICATIONS

toddergarten urban students: A study of early literacy instruction, treatment
quality, and treatment duration. Remedial and Special Education. Advance online
publication. doi: 10.1177/0741932510365359

FIELDS OF STUDY

Major Field: Education

Specialization: Special Education and Applied Behavior Analysis
TABLE OF CONTENTS

ABSTRACT.................................................................................................................. ii
Dedication...................................................................................................................iv
ACKNOWLEDGMENTS...............................................................................................v
VITA ..........................................................................................................................vi
LIST OF TABLES........................................................................................................xvi
LIST OF FIGURES.......................................................................................................xvii
CHAPTER 1 Introduction.............................................................................................1
Nationwide Reading Achievement............................................................................1
Teaching Reading .......................................................................................................3
  Good Instruction in the Primary Grades.................................................................3
Oral Reading Fluency.................................................................................................4
  Repeated Reading ...................................................................................................5
English Language Learners .........................................................................................8
  Good Instruction for ELLs.......................................................................................8
Availability of Technology .........................................................................................9
Read Naturally..........................................................................................................13
Purpose of the Study.................................................................................................15
Research Questions…………………………………………………………………………………..15

CHAPTER 2 Literature Review……………………………………………………………17

Introduction……………………………………………………………………………………..17

Reading Achievement across the Nation……………………………………………………..17

Importance of Effective Reading Instruction for Urban Learners…………………………21

Why it is Needed………………………………………………………………………..21

Placement in Special Education………………………………………………………22

Poor School and Post-School Outcomes………………………………………23

Five Components of Reading Instruction…………………………………………………24

Oral Reading Fluency………………………………………………………………………...24

Accuracy…………………………………………………………………………………25

Automaticity……………………………………………………………………………….25

Prosody…………………………………………………………………………………..26

Instruction in Oral Reading Fluency………………………………………………………27

Special Considerations…………………………………………………………………….27

Reading Level of Material……………………………………………………………..29

Passage Length………………………………………………………………………..30

Repeated Reading: An Instructional Approach to Teach ORF…………………………31

Providing Feedback……………………………………………………………………32

Providing Error Correction…………………………………………………………….34

Promoting Generalization………………………………………………………………35

ORF Interventions by Grade Level …………………………………………………………39
CHAPTER 3 Method.................................................................75

Participants .................................................................75

Selection Criteria .............................................................76

North River Academy .........................................................76

Tabitha .................................................................76

Diana .................................................................76

Javon .................................................................77

Angela .................................................................77

Quantasia .................................................................77

Tamika .................................................................78

Teachers .................................................................78

Experimenter and Secondary Observer .................................81

Setting .................................................................82

North River Academy .........................................................82

Definition and Measurement of the Dependent Variables ...............83

Primary Dependent Variables ........................................83

Correct Words per Minute ........................................83

Incorrect Words per Minute ........................................83

Comprehension Questions ........................................84

Oral Retell ..........................................................84

Secondary Dependent Variables .....................................84

DIBELS ..........................................................84

x
WJ-III ACH ................................................................. 85

Definition and Measurement of Independent Variable .................. 85

Key Words ........................................................................ 86

Cold Reading ..................................................................... 86

Read Along ........................................................................ 86

Practice Reading ............................................................... 87

Quiz Questions ................................................................... 87

Retell .................................................................................. 88

Pass Timing ......................................................................... 88

General Procedures ............................................................ 88

Experimental Design and Conditions ..................................... 88

Pretest ............................................................................... 90

Baseline ............................................................................. 91

Training .............................................................................. 92

Implementation of the Independent Variable - Read Naturally ....... 93

Treatment Probes .............................................................. 93

Generalization Probes ....................................................... 94

Posttest .............................................................................. 95

Materials ........................................................................... 95

Standardized Testing Materials ........................................... 95

Computers ......................................................................... 96

Read Naturally Software Edition .......................................... 96
Generalization Passages..........................................................97
Timer .....................................................................................97
Video Recorder.................................................................98
Counter...............................................................................98
Data Collection Sheets......................................................98
Reinforcers...........................................................................98
Interobserver Agreement ....................................................98
Procedural Integrity ............................................................99
Social Validity .................................................................100
CHAPTER 4 Results ............................................................102
Primary Dependent Variables ..............................................102
  Effect Sizes ......................................................................103
North River Academy.........................................................103
  Tabitha ...........................................................................104
  Diana ..............................................................................105
  Javon .............................................................................107
  Angela ...........................................................................108
  Quantasia ......................................................................109
  Tamika ..........................................................................110
Secondary Dependent Variables ........................................122
  Tabitha ...........................................................................122
  Diana ..............................................................................123
Javon .................................................................123
Angela .............................................................124
Quantasia..........................................................124
Tamika...............................................................125
Results: Social Validity.............................................129
Teachers ............................................................129
Participants.........................................................129
CHAPTER 5 Discussion...........................................131
Research Question One ........................................132
Correct Words per Minute.................................133
Incorrect Words per Minute ..............................134
Criterion Changes..............................................136
Research Question Two .......................................137
Research Questions Three .................................139
Research Question Four .................................141
Correct Words per Minute ..............................141
Incorrect Words per Minute...............................143
Criterion Changes..............................................144
Research Questions Five .................................145
Research Question Six .......................................147
Research Question Seven .................................148
Correct Words per Minute..............................148

xiii
Incorrect Words per Minute………………………………………...149
Criterion Changes…………………………………………………...149
Research Question Eight …………………………………….......................150
Research Question Nine…………………………………………………….151
Research Question Ten……………………………………………………..152
Research Question Eleven………………………………………………….153
Discussion of Secondary Dependent Variables…………………………….154
Limitations …………………………………………………………………156
Intervention Implemented in the Clinical Setting ………………….156
Comprehension Questions for Generalization Passages……………….157
Generalization in Classroom Reading Groups……………………...157
Limited Intervention Time………………………………………….158
Implications………………………………………………………………....158
Future Research …………………………………………………………….161
Summary…………………………………………………………………….163
REFERENCES ……………………………………………………………....166
APPENDICES ……………………………………………………………...182
Appendix A: Example RN Story and Comprehension Questions…..183
Appendix B: Example DI Story and Comprehension Questions…..186
Appendix C: Example AIMSweb Story and Comprehension Questions………………………………………………………….190
Appendix D: Experimenter Data Collection Sheet………………...193
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Popular computer programs</td>
<td>56</td>
</tr>
<tr>
<td>3.1 Participant demographic information and DIBELS winter benchmark scores</td>
<td>80</td>
</tr>
<tr>
<td>4.1 Baseline and treatment CWPM averages, including effect sizes and standard deviations</td>
<td>114</td>
</tr>
<tr>
<td>4.2 Criterion changes in CWPM</td>
<td>115</td>
</tr>
<tr>
<td>4.3 Baseline and treatment average words retold per minute</td>
<td>118</td>
</tr>
<tr>
<td>4.4 Baseline and treatment average percentage of comprehension questions answered correctly</td>
<td>121</td>
</tr>
<tr>
<td>4.5 Woodcock Johnson III pre- and post-test scores for LWID and WA subtests</td>
<td>126</td>
</tr>
<tr>
<td>4.6 Woodcock Johnson III pre- and post-test scores for RF and PC subtests</td>
<td>127</td>
</tr>
<tr>
<td>4.7 Winter and spring DIBELS ORF benchmark scores and risk status</td>
<td>128</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Treatment and generalization ORF data</td>
<td>112</td>
</tr>
<tr>
<td>4.2 Treatment and generalization retell data</td>
<td>116</td>
</tr>
<tr>
<td>4.3 Treatment and generalization comprehension data</td>
<td>119</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

For years we have known that many children start school below grade level; not having the skills needed to succeed. Even though children usually enter school ready to learn, for many children their enthusiasm for learning to read dwindles upon school entry (Mathes, Torgesen, & Allor, 2001). This could potentially exacerbate children’s reading problems, as Lonigan, Driscoll, Phillips, Cantor, Anthony, and Goldstein (2003) note that there is a strong link between children’s skills upon entering formal schooling and later academic success. Torgesen (1998) also mentions that children who have difficulty learning to read early on seldom catch up to their peers. To make matters worse, some students even have difficulty mastering phonological awareness skills that have been taught in basic reading instruction during the first few years of schooling (Macaruso, Hook, & McCabe, 2006), and these years are crucial for developing reading and writing in children (Teale, Paciga, & Hoffman, 2007). Thus, reading does not come naturally; it needs to be taught systematically and explicitly (Lyon, 1998; Torgesen, 1998).

Nationwide Reading Achievement

The Nation’s Report Card on reading (2007) reports that nationwide, students are progressing in reading. The percentage of students scoring at or above the basic level has increased, yet there was not a significant difference for students scoring at or above the proficient level. There is a somewhat parallel trend when disaggregating the data by race.
For example, the number of White, Black, Hispanic, and Asian/Pacific Islander students in the fourth grade who scored at or above basic was higher in 2007 than in 2005 and 1992. There were also higher percentages of White, Black, and Asian/Pacific Islander fourth graders who scored at or proficient, but neither of these changes were significant. What is worse is that only 14% of Black fourth graders scored at or above proficient, compared to 43% of White students scoring at or above proficient. Seventy-eight percent of White students and only 46% of Black students scored at or above the basic level.

These statistics on nationwide reading levels cause concern about the quality of reading instruction being provided in schools, particularly to minority students who live in urban districts. The NAEP also provides statistics on reading achievement of students from 11 urban districts across the nation, the Trial Urban District Assessment (TUDA, 2007). Overall, reading scores in most urban districts either remained stable or improved. Specifically, at eighth-grade, only two districts had higher percentages of students score at or above basic in 2007 as opposed to 2002; and two districts had higher percentages score at or above basic in comparison to 2005. Examining the data by district reveals that only 56% of eighth-grade students in Cleveland, Ohio schools scored at or above basic in 2007. Even though that percentage was higher than previous years, there were still 44% of eighth-grade students who cannot even read on a basic level. Trends for students with disabilities in urban schools are just as upsetting, with little change in average scores for students with disabilities in 2007. These data suggest that academic difficulty in reading persists for students across the nation. Students who do not read on a proficient level are likely to have poor school and post-school outcomes. Thus, it is vital that educators know the research on reading and implement effective practices in the classroom.
Many scholars would agree that children do not wake up one day and automatically know how to read. However, when it comes to reading instruction and what is needed for an individual to be a good reader, there has been long standing divisions among the academic community. On one side of the debate, teaching professionals argued for the use of the whole language approach, which is based on the notion that reading develops naturally and therefore does not emphasize the need for explicit instruction in phonemic awareness skills (Drecktrah & Chiang, 1997; Goodman, 1989). The other side believes that explicit phonemic awareness and systematic phonics instruction was the best way to teach children to read. Direct or explicit instruction involves an instructional paradigm that is teacher centered and emphasizes sequentially teaching reading skills that build upon each other (Carnine, Silbert, & Kame’enui, 1990; Drecktrah & Chiang).

*Good instruction in the primary grades.* Effective reading instruction must take place the day a child enters formal schooling. In order to prevent future reading problems, educators must take an early intervention approach to teaching reading to youngsters. Indeed, resources should be allocated for early identification and prevention of reading problems (Torgesen, 1998) as opposed to remediation. This intervention should take place by the end of second grade, as third grade students are no longer learning to read, but reading to learn. Therefore, it is essential that youngsters in preschool through second are taught how to read through explicit instruction.

Phonemic awareness and phonics are essential skills that need to be taught during early reading development (Lyon, 1998). Indeed, Lonigan and colleagues (2003) note the
importance of preschool children learning phonological skills as these skills are a prerequisite for decoding skills. Missall et al. (2007) concur in that learning foundational literacy skills early helps students develop other reading skills (i.e., decoding) more efficiently and effectively. Lyon also notes the importance of children learning how to automatically decode and recognize words in a fluent manner. Furthermore, these skills must be taught explicitly, as most children do not acquire these skills in the natural, print-rich environment. Indeed, explicit instruction has been noted as key for young children and children at-risk for reading failure (Lee & Vail, 2005; Lonigan et al., 2003; Lyon; Torgesen, 1998).

It is not enough to just teach the pre-reading and reading skills to children; they must become fluent in these skills. For example, Speece and Ritchey (2005) note that fluency instruction should probably begin and be combined with decoding instruction during the early stages of reading. Fluent reading is also highly predictive of later reading achievement (Cates & Rhymer, 2006); and necessary for later academic success overall, so educators must teach reading skills early in a child’s life (Lonigan et al., 2003).

Oral Reading Fluency

According to the National Reading Panel (NRP, 2000), oral reading fluency (ORF) is one of the five core components of reading instruction. Definitions of fluency are many, but most people agree that fluency is comprised of accuracy, automaticity, and prosody (Hudson, Lane, & Pullen, 2005; Kuhn & Stahl, 2003). Even though fluency is an essential part of being a proficient reader, many reading programs often neglect to include this component (Kame’enui & Simmons, 2001).
There is also a strong link between ORF and comprehension. This relationship is actually so strong that researchers (e.g., Fuchs, Fuchs, Hosp, & Jenkins, 2001) have noted that ORF is a better predictor of comprehension than more direct measures like question asking, retell, and cloze procedures. However, Therrien, Gormley, and Kubina (2006) point out that simply increasing a student’s ORF is not enough to ensure comprehension, even though it is an essential requirement for understanding text. Pikulski and Chard (2005) contend that, “…a deep construct views fluency far more broadly as part of a developmental process of building decoding skills that will form a bridge to reading comprehension and that will have a reciprocal, causal relationship with reading comprehension.” (p.511). Hence, ORF is a complex matter in which increasing a child’s fluency does not automatically result in increases in comprehension.

Nevertheless, there are researchers who support the idea that passage comprehension increases as readers become more fluent (Hasbrouck et al., 1999, Stevens, 2006). Stahl (2004) also cited numerous studies that have demonstrated that students who read connected text repeatedly or with assistance show increases in both reading comprehension and fluency measures. Kuhn & Stahl (2003) agree that fluency related instruction increases reading comprehension. One type of fluency related instruction that has the potential to impact ORF is repeated reading (RR).

Repeated Reading

Parker, Hasbrouck, and Denton (2002) declare that students sometimes need supplemental instruction even if the classroom teacher has an appropriate curriculum and quality instruction. Repeated reading is a viable option to increase the reading skills of these struggling readers. Interventions using repeated reading techniques have enjoyed
considerable success (Ardoin, Eckert, & Cole, 2008; Chafouleas et al., 2004; Chard et al., 2002; Cooke & Weinstein, 1992; Dowhower, 1987; Hapstak & Tracey, 2007; Hitchcock, Prater, & Dowrick, 2004; Lane et al., 2007; Martens et al., 2007; Vadasy et al., 2005; Welsch, 2007). Therrien and Kubina (2006) assert that RR techniques are appropriate for students whose instructional reading level is between first and third grade, because these students have already acquired a basic foundation in reading (i.e., blending). Furthermore, the authors contend that RR is also beneficial for students reading above the third grade level, but read slowly and haltingly.

Repeated reading in its truest form is when students are given short passages to read either to a predetermined criterion level and/or a specific number of times, and then they are given a new passage to read (Mastropieri et al., 1999; Samuels, 1979). Repeated reading has evolved over the years though, and researchers have begun adding other components to repeated reading interventions. One component that can be included in RR interventions is a listening passage preview (LPP). This is often beneficial to dysfluent readers because fluent reading is modeling in the form of an adult reading the passage while the student follows along.

Daly and Martens (1994) studied LPP as part of an intervention used to increase the ORF of four boys with learning disabilities. The researchers found the highest increases in fluency and accuracy performance in the LPP condition, although there were more dramatic effects for two of the four participants. The authors purport that LPP contained a generalization component because the student reader heard unknown words from the classroom reading curriculum modeled during this condition. Therefore, it was
expected that the students would have the greatest increases in fluency and accuracy during this condition.

Welsch (2007) also compared LPP and RR approaches for increasing the ORF of four boys with learning disabilities. He found that ORF increased the most during the RR condition. Initially, however, the generalization of ORF skills was limited for 3 of the 4 participants, so Welsch decided to use a changing criterion tactic in order to boost generality. He calculated the average number of CWPM in the previous four generalization passages and increased that amount by ten percent. Once the students met that criterion for three out of four sessions, he increased the criterion an additional ten percent. This changing criterion tactic resulted in an increase in CWPM, which was proportionate to responding on the RR treatment passages.

Bonfiglio, Daly, Martens, Lin, and Corsaut (2004) also individualized CWPM goals for their participant in a performance-based treatment condition. Here the goals were changed everyday, and were based on the participant’s highest number of CWPM on any passage leading up to that point. Her performance increased above baseline levels in this and other treatment conditions, noting the importance of goal setting. Setting performance goals seems like a logical topic for future research as a meta-analysis on RR by Therrien (2004) found large effect sizes for fluency (1.70) when the interventions used a performance criterion to decide when to move onto the next reading passage. Providing a corrective feedback component with the performance criterion also promoted overall gains in ORF and comprehension. Through this meta-analysis, he determined that RR can improve ORF and comprehension in untrained (generalization) passages.
English Language Learners (ELLs)

American schools are becoming considerably more diverse. Teale (2009) points out that there are about 5 million ELLs in schools throughout the United States. Furthermore, ELLs are no longer mainly in states like California, Texas, Florida, and New York; but are expanding into states like Minnesota, Alabama, North Carolina, and Indiana. Teale also states that ELLs tend to be educated in urban schools, so teachers need to learn how to provide the best literacy instruction possible.

*Good Instruction for ELLs*

Gersten and Baker (2000) discuss effective instruction for ELLs. Their report coincides with the opinions and research of others (Denton et al., 2004; Huebner, 2009; Linan-Thompson, Cirino, & Vaughn, 2007; Linan-Thompson, Vaughn, Hickman-Davis, & Kouzekanani, 2003; Pikulski & Chard, 2005; Teale, 2009) These researchers note the importance of providing ELLs with evidence-based literacy instruction just like one would provide any student. Educators need to focus on providing instruction that keeps the ELLs actively engaged; provides multiple opportunities for practice; and incorporates conversation skills. Additionally, the teachers should monitor and assess student progress on an ongoing basis. The teachers should provide instruction in phonemic awareness, phonics, fluency, vocabulary, and comprehension. If an ELL has difficulty in any of these areas, teachers should provide small group instruction to learners who exhibit the same needs (Huebner). Gersten and Baker purport that any good English-language development program include instruction in: a) developing proficient and fluent English; b) formal grammar; and c) learning academic content.
Many children, especially urban learners and ELLs, could benefit from research-based reading interventions. With so many students evidencing a need for supplemental reading instruction, and not enough teachers to provide it, computers might begin to play an important role in reading interventions. Technology cannot replace teachers, but when available, it sure can be of great assistance.

Availability of Technology

The use of technology has become quite commonplace in everyday life. Many people have access to technology on the job, in school, and even in their homes. Nowhere is this truer than in the suburbs. Children who live in suburban neighborhoods are often afforded greater opportunities than those who live in urban or rural areas. The quality of suburban schools also tends to be better than that of urban and rural schools. For example, Blanchett, Mumford, and Beachum (2005) point out that suburban schools are often deemed “high performance” schools. The students in these schools are often educated in newer classroom buildings with state-of-the-art technology labs, staffed with qualified teachers that are paid better than teachers in urban areas, and offer honors and advanced placement classes. Students in these suburban schools are afforded these advantages due to the inequalities that exist between wealthy (often suburban) and poor (urban or rural) school districts (Gardner & Miranda, 2001). Researchers have speculated that high-income families actually help their school districts by raising more money (external funds) for computers (Hess & Leal, 2001), which would provide students in suburban schools even greater access to technology. Hess and Leal found that higher-income urban districts provide more access to technology that decreases the student-to-computer ratio than low income urban districts. One might generalize this information to
suburban districts in that these schools generally have more money per capita, therefore, extra money might be spent on school improvement projects, like providing more technology.

Urban areas, on the other hand, present unique challenges relative to providing optimum learning environments for all students. Lack of resources, interventions, and professional development opportunities are among some of the more salient impediments that must be overcome in urban schools in order to promote equity. However, teacher quality is another major hindrance to students in urban schools being provided with the best education possible. It is not uncommon for inexperienced teachers to be placed in schools where students exhibit the greatest need. In fact, Gehrke and McCoy (2007) note the difficulty in recruiting and retaining qualified special education teachers in “hard-to-staff” schools (i.e., urban, rural, poor).

A lack of highly qualified staff exacerbates an already downtrodden environment in urban schools. There is a broad spectrum of disparities that exist among urban schools, from the distribution of technology (Raynor, 2006; Hess & Leal, 1999; Hess & Leal, 2001) to the availability of other resources, including appropriate [and enough] curricular materials (Raynor; Lalas, 2007; Swanson et al., 2007). From a social justice perspective, Lalas states that these inequities must first be recognized in order to be diminished, and that resources need to be allocated in an equitable manner. Nevertheless, even when educational reform is taking place, care must be taken to explicitly state how resources will be distributed to schools in a more equitable fashion, so that “systemic barriers to excellence” can be checked (Raynor, p.54). To be sure, educational equity in urban districts cannot be achieved through just technology grants and special funding, schools
have to actively search for technology partnerships and redistribute existing financial resources for technology (Walker, 1997).

Hess and Leal (1999) agree that technological resources are of great importance to education in America, especially for students in urban settings. Their assertion is that access to technology should be equitable and substantial, so as to not create informational “haves” and “have-nots.” In fact, a high school student from New Orleans mentions that two or more students often had to share one computer at the same time (Raynor, 2006). Similarly, at another large urban high school, nearly 4,000 students had to share 16 computers in the library (Brown, 2006). Brown mentions the indispensable nature of accessing this technology in the school library, because a lot of students from low socioeconomic households do not have computers and/or internet access at home.

To be sure, having access to computers provides a stepping stone for proficient use of computers and technology (Hess & Leal, 2001). However, the ratio of students-to-computers in urban districts is not as good as suburban districts. According to Hess and Leal, there is a “digital divide” between types of districts (i.e., suburban, urban), but the variables that affect each are not the same, so these systems cannot be equally compared. What is important, they note, is how computers are distributed across urban school districts to low-income and minority to students who already have serious educational obstacles to overcome. The authors examined the most recent computer provision data from the Council of Urban Boards of Education to determine the correlates of the likelihood of computer access across urban school districts. Results revealed that urban districts with more African American students have a higher student-to-computer ratio. But, the authors do note that the racial digital divide has fallen by approximately half
since their last study in 1999. An additional finding was that urban districts with more resources, and per-pupil spending, had a higher provision of classroom technology than those with less district funds. Hess and Leal refer to this as the “interurban digital divide.” Thus, urban districts who receive less funding provide their students with less computer instruction. The interurban digital divide is evidence that inequities do not just exist across different types of school districts, but also across urban districts. Hess and Leal maintain that such inequities between urban communities will actually reinforce or exacerbate these pre-existing gaps.

One at-risk, poor urban high school experiencing steady declines in attendance actually phased in computer technology over the course of seven years. Diem and Katims (2002) conducted a longitudinal study to determine the effects of this new computer technology on the students, teachers, and administrators. Over the course of the seven years, test scores on standardized state tests showed increases across reading, writing, and math, but the school’s scores were still the third lowest out of eight other high schools in the same district. Yet and still, the teachers, students, and administrators noted positive results overall.

The above research supports the use of technology to enhance academic instruction. However, the resources must be available to buy computers and educational software. Computer-assisted instruction is one way of providing individualized supplemental instruction for struggling learners. There are various computer programs that have been developed to boost academic performance. One such program that might foster academic success in reading is Read Naturally.
Read Naturally

*Read Naturally* (RN) is a program that uses scientifically-validated techniques to enhance the ORF skills of children in first through eighth grade. In fact, Welsch (2006) includes some of the components that the RN program employs as ways to increase ORF. Some of the strategies he discusses are: RR with a model; using read-along/paired reading; and using computers. It is important to point out that there are two versions of the RN sequenced series: a Master Edition (ME), which uses audiotapes, CDs, and paper versions of the story; and a Software Edition (SE), which uses computer technology. The few published studies on the effectiveness of the RN program have been used with RN (ME).

Hasbrouck, Ihnot, and Rogers (1999) reported the effects of the RN (ME) program on remedial students in elementary school and special education students in middle school. After the RN (ME) program was developed, it was tested in various classrooms in different states, with various grade level students. Hasbrouck and colleagues reported the results for 214 second and third grade Title 1 students. These students used RN (ME) 25 minutes daily for an average of 32 weeks. Their ORF scores were quite low at the beginning of each school year (below the 25th percentile), but all of the students’ scores moved to between the 25th and 50th percentile in the spring.

Denton, Anthony, Parker, and Hasbrouck (2004) used *Read Well* and RN (ME) to determine their effectiveness on the reading skills of 93 bilingual students in second through fifth grade. Tutoring occurred three times a week for 40 minutes over the course of 10 weeks. Results revealed that the *Read Well* program significantly increased the tutored students’ decoding, but there were no significant differences between groups on
comprehension measures. The students in the RN program made negligible gains in comprehension and little to no gains in decoding. However, the authors speculate that the lack of comprehension effects might be due to an increased focus on ORF in the RN program, and for ELLs who are not reading in their native language, more time may be needed to understand what they have read.

A later study (Denton, Fletcher, Anthony, & Francis, 2006) used RN (ME) with students who had persistent reading difficulty. In fact, the five first grade participants were repeating that grade and had not responded to previous intervention efforts. A major difference between this and other RN studies is that Denton and colleagues used another intervention program, Phono-Graphix, for eight weeks prior to intervening with the RN program. The two interventions focused on different skills though; with a decoding emphasis in Phono-Graphix and an emphasis on fluency in RN. The students received the RN intervention for an hour a day for eight weeks following the Phono-Graphix intervention. Results revealed that the RN program produced effect sizes for fluency that were small to moderate. The struggling readers in this study made gains in ORF of word lists and connected text; however, gains in comprehension were negligible.

Gibson (2009) actually used RN (SE) with eight first grade students who were at risk for reading failure. There were two intervention phases. During phase one, the participants had a goal criterion of 40 CWPM on RN probes. During phase two, each participant’s goal was increased based on their individual performance during phase one. This was done in an effort to promote fluency gains on generalization passages. The researcher found that the RN software was effective at increasing the ORF and comprehension of all eight participants, generalization gains were only demonstrated
after implementing phase two, and 5 of the 8 participants were able to reduce their risk status according to DIBELS.

Purpose of the study

This study investigated the effects of a computer-assisted reading program, Read Naturally (RN) on the ORF, comprehension, and generalization skills of urban second grade students who were at-risk for or had a disability. Additionally, the goal criterion of CWPM was systematically changed to determine its effect on generalization. This was done to extend the research done by Gibson (2009).

To date, there are limited studies regarding the effectiveness of the RN program (Denton et al., 2004; Denton et al., 2006; Hasbrouck et al., 1999), and in particular, the software edition (Gibson, 2009). Therefore, this study examined the effects of the RN (SE), on the ORF, comprehension, and generalization of second grade students at risk for reading failure. Specifically, the following research questions were addressed.

Research Questions

1) What effect will the RN computer software program have on the oral reading fluency of urban second-grade students at risk for reading failure?

2) What effect will the RN computer software program have on the number of comprehension questions answered correctly by the participants?

3) What effect will the RN computer software program have on the oral retell of the participants?

4) What effect will the RN computer software program have on the oral reading fluency on generalization passages (i.e., AIMSweb) of the participants?
5) What effect will the RN computer software program have on the number of comprehension questions answered correctly on generalization passages (i.e., AIMSweb) by the participants?

6) What effect will the RN computer software program have on the oral retell on generalization passages (i.e., AIMSweb) by the participants?

7) What effect will the RN computer software program have on the oral reading fluency of generalization passages conducted in the classroom using the classroom curriculum?

8) What effect will the RN computer software program have on the number of comprehension questions answered correctly on generalization passages conducted in the classroom using the classroom curriculum?

9) What effect will the RN computer software program have on the oral retell on generalization passages conducted in the classroom using the classroom curriculum?

10) How will the participants respond to the RN program (as measured by a social validity questionnaire)?

11) How will the teacher rate the effectiveness of the RN intervention on the ORF and comprehension of the students?
This chapter contains a review of literature on oral reading fluency and computer-assisted instruction (CAI) in reading. The ORF review focuses on some special considerations when teaching fluency, strategies to teach ORF, increasing generalization, and teaching ELLs. The literature on CAI in reading reports on the benefits of technology, various computer programs used to increase reading skills for children in the primary grades, and current research of CAI in reading.

Introduction

Reading Achievement across the Nation

Reading failure has a major impact on a person’s life. There are approximately 23 million adults who function on or below a basic level, rendering them functionally illiterate (Hall, Hughes, & Filbert, 2000). The current statistics on reading across the nation are not especially encouraging. According to The Nation’s Report Card on reading (2007), students are making gains in reading across the nation at or above the basic level; however, there was not a significant difference for students scoring at or above the proficient level. Scoring at the basic level means only partial mastery of skills needed to work proficiently on grade level, whereas the proficient level denotes solid understanding.
Analysis of The Nation’s Report Card on reading (2007) reveals that there was no significant difference in average scores for eighth graders in 2007 as opposed to in 1992 and 2005. In fact, eighth graders only scored one point higher in 2007 than in 2005; and there was no significant change in the percentage of eighth graders who scored at or above proficient on the 2007 reading assessment. Seventy-four percent of eighth-graders across the nation scored at or above the basic level, but only 31% of those students scored at or above the proficient level. This means that almost three-quarters of eighth grade students are not reading on a proficient level and are entering high school unprepared for the academic tasks that lie ahead.

The Trial Urban District Assessment (TUDA, 2007) is another test in which the NAEP gathers statistics on the reading achievement of students from 11 urban districts across the nation. These urban districts provide education to a higher percentage of minority and lower-income households than the in the nation as a whole. However, scores on the TUDA do not provide sufficient evidence that these urban schools are educating minority and low-income students to the best extent possible. For instance, when comparing only low-income students in the urban schools to other low-income students in the nation, seven of the urban districts scored lower than the nation at grade 4; and five scored lower at grade 8. Relative to racial groups, 2007 average reading scores in only one of the urban districts were higher for all three racial groups in the fourth grade; while eighth-grade scores were only higher for Black students in one district and Hispanic students in a different district. In Cleveland, Ohio, for example, Black students did not make any significant changes in average scores in fourth or eighth grade, with
fourth-grade students actually scoring one point less in 2007 than in 2005. The results for Hispanic students in Cleveland parallel those of Black students.

Disaggregating the scores by gender and racial group were also revealing. Across the nation, females outperformed males at both grades 4 and 8; however, the only increases in scores since 2005 were for males and females in grade 4, not grade eight. Females of all races outperformed their male counterparts in both grades as well; with the greatest average gender gap occurring in the eighth grade. One will note, then, that the gap widens as students progress in school. Black males obtained the lowest average scores for race and gender in both grades, providing evidence for the need to use evidence-based reading instruction for prevention and remediation of reading difficulty for all students, especially African-American boys.

These results from the Nation’s Report Card and TUDA provide evidence of an achievement gap, which exists when children from certain backgrounds score significantly lower on standardized test instruments than children from other backgrounds. For example, even though White, Black, and Hispanic students made gains in reading, these gains did not always serve to narrow the achievement gaps with White students. In the 2007 assessment, there was no significant change in the White-Black gap at grade 8; however, there was a decrease in this same gap at grade 4 (as opposed to the gaps in 1992 and 2005). Results from the TUDA also reveal that there was little change in the performance gaps by race. None of the participating urban districts showed a significant change in gaps between Black and White students in 2007 at grade 4, while the gap between Black and White students at grade 8 was actually larger in Los Angeles.
than in prior years (2002 and 2005). These statistics are disconcerting as the goal is to narrow the achievement gap between races.

Teale and colleagues (2007) point out that achievement gaps exists for impoverished children in reading and writing as compared to children from middle and high-income children; and another gap exists for African American and Latino children compared to Caucasians who score higher. Some researchers purport that the gap between minority and nonminority students actually widens as students progress from first grade up through third and even fourth grade (Shippen, Houchins, Calhoon, Furlow, & Sartor, 2006). The achievement gap for African Americans is further widened by disciplinary exclusion, creating a “domino effect” in which these excluded students receive ineffective, remedial instruction that does little to increase their academic achievement and increases the likelihood of being retained (Townsend, 2000).

Teale et al. (2007) assert that in order to narrow the achievement gap, educators must bridge the curriculum gap that exists in urban schools. This curriculum gap exists when teachers either do not teach, or scarcely teach some of the essential components that successful readers and writers need. The three areas that the authors note are lacking in urban classrooms are instruction in comprehension, writing, and content areas such as science and social studies – developing background and domain specific knowledge. Belfiore, Auld, and Lee (2005) offer that to close the achievement gap, urban schools must provide students with plenty of opportunities to become academically successful early and often in their schooling. Furthermore, the authors propose that the achievement gap will not be closed unless teachers: believe all students can learn; design and implement sound pedagogy; and create academically challenging curricula. The
curriculum could also include culturally-relevant content so that students feel appreciated by their teachers (Walker-Dalhouse & Risko, 2008). These authors’ opinions are based on research conducted in schools and in national reports.

Importance of Effective Reading Instruction for Urban Learners

Effective, systematic reading instruction is paramount in urban schools, as these schools tend to be populated by poor minority students. The effects of poverty can be quite devastating to a person’s overall life, and this is especially true when it comes to education. The possibility of being at-risk for school failure and/or being placed in special education is impacted both directly and indirectly by living in poverty (Artiles, Harry, Reschly, and Chinn, 2002). In addition, Skiba et al. (2008) also note that minorities are at risk of being referred for special education due to risk factors that are associated with poverty.

Why it is needed. Many urban learners come to school lacking the knowledge and background experiences necessary for proper reading development; therefore, it is of the utmost importance that these learners receive solid instruction in the foundational reading skills. Urban learners tend to start school with a disadvantage related to reading, primarily, fewer language skills (e.g., smaller vocabulary, lack of phonemic awareness, etc.), so it is imperative that teachers provide interventions that increase the pre-reading and reading ability of children in the primary grades, especially in fluency.

Children also come to school at-risk for reading failure if they live in poverty, have limited English proficiency, if their parents have low reading levels, or if the children have disabilities in speech, language, or hearing (Lyon, 1998). The addition of poverty to minority status further exacerbates academic difficulty. Often minority and
low socio-economic status (SES) children enter formal schooling with deficits in phonemic awareness and general oral language skills (i.e., vocabulary), so these children will need specific interventions in order to develop these skills and make sufficient reading progress (Torgesen, 1998). Sufficient reading progress is essential if the achievement gap between poor, minority students and their Caucasian counterparts is to be narrowed. This issue is especially important in urban schools as these schools tend to overwhelmingly educate impoverished African American and Latino students (Teale et al., 2007).

Placement in Special Education. African American children, particularly those who are poor, are at a greater risk of being placed in special education. Even when poverty is held constant, race has a significant effect on whether or not a student is placed in special education (Skiba, Poloni-Staudinger, Simmons, Feggins-Azziz, & Chung, 2005). This is a problem as it is quite difficult for these same students to actually exit special education, or even be served in the least restrictive environment. Furthermore, Krezmien, Leone, & Achilles (2006) found higher rates of suspension in students with disabilities as opposed to those without. Additionally, African American students with disabilities were more likely to be suspended than any other of the racial groups with disabilities. So, African American children in special education are excluded to a greater extent. In turn, these students fall further behind academically and are not offered any behavior intervention or support. Accordingly, students who experience persistent poor academic performance will be at-risk for special education (Gardner & Miranda, 2001). Thus, it is the culture and organization of schools that actually increase the risk that minorities will be placed in special education (O’Connor & Fernandez, 2006).
Without a doubt, these urban students need to be taught with research-based programs that will remedy the reading problems that they might experience (Shippen et al., 2006). This is especially the case for minority students who come to school with considerably fewer readiness skills than their peers. Schools must have a strong curriculum that focuses on these skills (Cartledge, 2005), and should teach all of the foundational skills needed to promote lasting literacy growth (Teale et al., 2007). Belfiore et al. (2005) assert that students in poor-urban schools need to want to engage in academics as opposed to non-academic behavior, and in order for this to happen, schools must provide curricula that are challenging and supportive. If not, the effects will most likely be academic failure.

*Poor school and post-school outcomes.* Difficulty with reading poses problems across curricular areas, over the child’s whole academic career, and even post-school life. For instance, Simner and Barnes (1991) reported that students who perform poorly in reading and math in first grade are actually at an increased risk to drop out of high school than students who earned A’s and B’s. Students with disabilities also drop out of school at higher rates than their nondisabled counterparts (Drakeford & Staples, 2006).

Oftentimes, students who struggle with reading will exhibit an externalizing behavior that will remove them from the aversive situation, frequently resulting in suspension. Whether suspension is in or out-of-school, the child does not have access to reading instruction, therefore no learning can take place. Arcia (2006) found an association among higher suspensions, poor reading, and higher school dropout rates. This is consistent with Townsend (2000) who proposed a relationship with frequent school exclusions, illegal activities and leaving school early. An obvious implication is a
need for effective reading interventions for urban learners showing risk for reading failure.

Five Components of Reading Instruction

In 1997, the National Institute of Child Health and Human Development (NICHD) convened a national panel to assess the effectiveness of different approaches used to teach children to read. For over two years, the National Reading Panel (NRP, 2000) reviewed research-based knowledge on reading instruction and identified five essential components of reading which include phonemic awareness, phonics, fluency, vocabulary, and comprehension.

Instruction in phonemic awareness focuses on analyzing and manipulating the sounds (phonemes) of words and blending these sounds to say complete words (Ehri, 2004). Phonics is the understanding that there is a predictable relationship between phonemes and graphemes, and children are taught how to use these associations to read. Vocabulary at its most basic level is understanding the meanings of words, and it should be taught directly and indirectly (Kamil, 2004). Comprehension is the ability to understand printed text and is vital to the development of children's reading skills and their ability to obtain an education (NRP, 2000). The fifth component of balanced reading instruction and the hallmark of a good reader is being able to read fluently, with understanding.

Oral Reading Fluency

How to teach children to read has been a topic of debates for years. Fluency instruction is no different. Even though instruction in reading fluency is vital for developing proficient readers, reading programs often neglect to include this component
It may be that researchers disagree on the definition of fluency and the most effective interventions to teach it, but most authorities would agree that fluency is a necessary component of proficient reading (Chard et al., 2002; Hasbrouck et al., 1999; Hudson et al., 2005; Kuhn & Stahl, 2003).

So what exactly is oral reading fluency (ORF)? Some related answers might be “reading rate,” “reading with expression,” “reading fast,” “speed or ease of reading” and “recognizing words automatically.” In general, “automaticity” and “fluency” have been used interchangeably (NRP, 2000) over the years. Since there is no concrete definition of fluency, people are left to define the term however they choose. Nevertheless, there are three components of fluency upon which most people would agree: accuracy, automaticity, and prosody (Hudson et al., 2005; Kuhn & Stahl, 2003).

Accuracy: One of the components that is essential for fluent reading is accuracy (Kuhn & Stahl, 2003). Reading words accurately is simply “the ability to recognize or decode words correctly” (Hudson et al., 2005, p.703), but decoding can be slow for developing readers who exert more energy sounding out words (NRP, 2000). Stahl (2004) points out that a strong knowledge of the alphabetic principle increases the likelihood that children will read words accurately. The author also notes that students who lack decoding skills will have difficulty reading words accurately in both word lists and connected text. This becomes especially problematic since the words that students tend to misread are usually essential to gathering meaning from the text, thereby decreasing the child’s comprehension level (Hudson et al.). In fact, reading fluently in connected text seems to have a greater impact on comprehension levels than does fluent reading of word lists (Fuchs et al., 2001).
**Automaticity.** Automaticity, which refers to reading words quickly and effortlessly (Hudson et al., 2005; Stevens, 2006) is another component that is essential for fluent reading (Kuhn & Stahl, 2003). The term “rate” has sometimes been used synonymously with automaticity. This is because oftentimes fluent reading is measured in the number of correct words read per minute (Hudson et al.). Sufficient decoding skills are essential in order to read words automatically.

Reading words accurately goes hand-in-hand with automatic recognition of words (Hudson et al., 2005, Kuhn & Stahl, 2003). As readers become proficient with decoding they begin to read words automatically and faster, chunking words into phrases (Stevens, 2006), freeing cognitive effort (Hudson et al.), which generally leads to increased comprehension. Similarly, Kuhn and Stahl mention that a lack of decoding skills detracts from the ability to gather meaning from text. Moreover, Chard et al. (2002) mention that students with learning or reading disabilities tend to struggle with decoding and automaticity. Therefore, it is imperative that readers, especially those with disabilities, are taught how to recognize words automatically, which can be achieved by providing children with ample opportunities to practice these skills. The NRP (2000) also underscores the importance of practice noting that automaticity of word reading usually requires an extensive amount of training.

**Prosody.** Prosody is probably the most neglected component of ORF in research literature as well as in the classroom. In general, too few teachers systematically program for prosodic reading. Reading with expression is the phrase that most people use to refer to prosody. More concretely, the term prosody refers to the variations of intonation, pitch stress patterns, timing, and loudness that all contribute to reading text with expression
(Hudson et al., 2005; Kuhn & Stahl, 2003; Stevens, 2006). Reading with expression involves paying attention to other features of the text, like punctuation, which can “signal question, surprise, exclamation, and other meanings beyond the semantics of the words being spoken” (Hudson et al., p.704).

In this way, prosodic reading should sound as if the reader is speaking aloud, or having a conversation with someone. There should be variations in the speed, the tone, and the length of words the reader says. Nonetheless, dysfluent readers tend to read in a more monotone voice and combine words into meaningless phrases or sentences (Hasbrouck et al., 1999; Hudson et al., 2005; Rasinski, 2006). Appropriate chunking of words into meaningful phrases or groups is essential for understanding text (Kuhn & Stahl, 2003), although printed text does not offer much guidance as to appropriate phrasing or prosodic reading as do conversations (LeVasseur, Macaruso, & Shankweiler, 2008). Consequently, the role of accuracy, automaticity, and prosody taken together account for the need to explicitly teach fluency skills to developing readers. As per Miller and Schwanenflugel (2008), definitions of fluency should include prosody as a main component.

**Instruction in Oral Reading Fluency**

*Special Considerations*

As ORF is one of the key components needed for proper reading development, learning to read fluently cannot be left to chance. Many students with disabilities do not read on a proficient level (Welsch, 2007), and problems with fluency, in particular, seems to be a common attribute of students with disabilities (Mastropieri et al., 1999). Likewise, fluency instruction is essential for urban learners and boys in particular, as the Nation’s
Report Card (2007) points out that girls outperform boys on reading assessments. However, before one begins teaching ORF, there are some special considerations that must be taken into account.

There are various components or subskills that the umbrella term “fluency” comprises, that are often left out of definitions and research regarding fluency (Wolf & Katzir-Cohen, 2001). These include letter naming, letter sound naming, nonsense word fluency (NWF), sentence fluency, and rapid automatized naming (RAN), to name a few (Speece & Ritchey, 2005, Wolf & Katzir-Cohen). Hudson and colleagues (2005) also note the importance of subprocesses in decoding words accurately as a precursor to fluent reading. Likewise, Speece, Mills, Ritchey, and Hillman (2003) suggest that letter fluency tasks which involve children naming letters or their sounds accurately and automatically, could impact their subsequent competence in reading. In fact, Speece and Ritchey (2005) suggest that sublexical skills, such as phonological awareness, can be used to predict levels of ORF as early as first grade. In short, “reading fluency involves every process and subskill involved in reading.” (Wolf & Katzir-Cohen, p.220) which is not shocking considering that ORF is a complex, multifaceted skill (Fuchs et al., 2001).

Because of the complexity of ORF, it is no surprise that many of our school-age population continue to struggle to perform this skill (NRP, 2000). Readers who are not fluent tend to read laboriously, often saying one word at a time, ignore punctuation, and read with limited expression (Hasbrouck et al., 1999; Hudson et al., 2005). This type of reading severely limits comprehension by limiting the amount of working memory a student has to think about what they are reading (Chard et al., 2002; Mastropieri et al., 1999; NRP). In addition, students who lack fluency skills might feel like a failure and
may not want to participate in reading activities in class or for leisure (Hasbrouck et al.; Rasinski, 2006). Since fluent reading might be considered the cornerstone of proficiency and has such negative consequences if not achieved, it is imperative that students are explicitly taught how to read more accurately, automatically, and prosodically.

Before teaching fluency, educators need to realize that there are prerequisite skills that students need to be proficient in to facilitate fluent reading. Among them are phonemic awareness; letter and word fluency; and decoding. For instance, students who have not acquired proper phonemic awareness skills find it difficult to learn to read words because they have not mastered letter recognition or being able to sort pictures based on initial sound (Johnston, 2000). Torgesen (1998) asserts that measures of word reading fluency are the most important indicator of progress that children in early elementary school will make in reading. Therefore, decoding skills are important for fluent reading. Children need to be able to recognize words in order to read text fluently (Lewandowski, Begeny, & Rogers, 2006). Students with the most severe reading problems might also need to attain a particular level of decoding competency in order to benefit from instruction in ORF (Denton et al., 2006).

Reading level of material. There is also some debate as far as the level of reading material that should be used when teaching and practicing ORF. Spear-Swerling (2006) points out that researchers have traditionally used grade-appropriate texts to measure ORF, which might be frustrating for poor readers. However, by using easier reading materials, fluency levels could be measured more easily because all students would be able to read the material accurately, leaving only speed of reading as the differentiating factor. Similarly, Gibb and Wilder (2002) mention that reading instructional or
independent level passages tend to increase students’ ORF. O’Connor, Bell, Harty, Larkin, Sackor, and Zigmond (2002) studied the effects of reading level material versus grade level material (i.e., frustration) on ORF of students in the intermediate grades, and found that students who were instructed using material on their reading level, and started out with a lower fluency level actually made greater gains than those instructed in grade level material. Likewise, Daly, Martens, Kilmer, and Massie (1996) noted that students improved their ORF when assessment materials matched the student’s skill level and were similar to instructional materials. Sindelar, Monda, and O’Shea (1990) used repeated readings (RR) as an intervention with students in the intermediate grades and reported that mastery level students outperformed instructional level students on ORF.

On the other hand, Kuhn and Stahl (2003) reviewed fluency related literature and hypothesized that more difficult reading material would actually improve ORF. Readability formulas must be used in order to determine the level of the reading passage being used to teach ORF, but Ardoin, McCall, and Klubnik (2007) cite research that suggests that readability formulas do not predict passage difficulty satisfactorily. Therefore, it behooves researcher to verify the level of reading material by using at least two different readability formulas.

*Passage length.* Another factor that must be taken into consideration when teaching ORF is the length of the passages. Repeated reading is one effective method that has been used to teach ORF. With this method, students are generally given short passages to read either to a predetermined criterion level and/or a specific number of times, before they are given a new passage to read (Mastropieri et al., 1999; Samuels, 1979). However, the lengths of the passages may vary according to the student’s reading
level and difficulty of the passage. In general, the length of the passage should be around 100-150 words. Researchers have consistently developed passages within this range, as students in the Eckert, Dunn, and Ardoin (2006) study read passages containing between 81-144 words, with an average of 105.25; and students in the Begeny, Daly, and Valleley (2006) study read passages with between 110-140 words per passage. As with any instructional method, it makes sense to develop your materials in order to support the individual needs of learners.

Educators must be well aware of the complexity of reading fluency. There are subskills (i.e., letter naming) and prerequisite (i.e., decoding) skills of fluency that need to be taken into consideration before planning fluency instruction. Additionally, the difficulty level and length of the reading material must be taken into account. Once educators have sufficiently planned for these variables, instruction in reading fluency can begin.

Repeated Reading: An Instructional Approach to Teach ORF

One method that has been demonstrated to increase ORF is repeated reading, which Samuels (1979) notes should accompany existing reading instruction. Repeated reading can be used to improve skills such as word recognition, comprehension, reading rate, and expression (Johnston, 2000). Students are generally given short passages to read either to a predetermined criterion level and/or a specific number of times, before they are given a new passage to read (Mastropieri et al., 1999; Samuels, 1979). Chafouleas et al. (2004) note that this type of instruction is beneficial for students to increase their ORF since it involves drill and practice. Repeated reading can involve reading words/lists repeatedly (e.g., LeVasseur et al., 2008; Vadasy et al., 2005) or repeatedly reading
connected text (e.g., Hitchcock et al., 2004; LeVasseur et al.). Over the years, various techniques have been developed that are modified forms of repeated reading. These include: independent or unassisted reading; assisted, choral, partner or unison reading; and echo reading (Hitchcock et al.; Kuhn & Stahl, 2003; Vadasy et al.). The aforementioned authors define the various terms in much the same manner. Unassisted reading refers to the student reading text independently. On the other hand, assisted reading involves the student reading simultaneously with another person (i.e., teacher, peer, paraprofessional); while echo reading requires the skilled reader (or tutor) to read a phrase or sentence aloud followed by the student reading the exact phrase or sentence.

Interventions using repeating reading techniques have enjoyed considerable success. In a literature review by Chard et al. (2002), repeated reading was demonstrated to be effective at increasing the reading rate and accuracy, as well as comprehension of elementary students with learning disabilities (LD). Likewise, studies that used repeated reading as part of an intervention implemented with first grade students have demonstrated increases in ORF (Hitchcock et al., 2004; Lane et al., 2007; Vadasy et al., 2005). In a study by Hapstak and Tracey (2007), assisted repeated reading was used as the sole intervention with four first-grade students of varying reading ability. All four students increased their overall words read correct per minute (CWPM) from baseline to intervention and from the first reading each session to the final reading of the session.

Providing feedback. A skill must be taught to mastery in order for performance to be considered fluent. Educators should ensure that students learn the target skill without making too many mistakes, and if mistakes should occur, there should be a process to remediate such errors. One method that might provide an additional boost in performance
is to give the student feedback regarding their performance during reading fluency tasks. Types of performance feedback might include public posting, self-scoring, response cards, or verbal feedback (Eckert et al., 2006).

Eckert and colleagues (2006) note that until recently, there has been little examination of the effects of performance feedback on ORF skills. Therefore, the authors conducted a study with six second-grade general education students who were having difficulty with ORF, but were not receiving any special education or intervention services. There were two experimental conditions: performance feedback on words read correctly (PFWC) and performance feedback on words read incorrectly (PFWI). The students received performance feedback bar graphs indicating the number of words read correctly or incorrectly depending upon which experimental condition was in effect at that time. Sessions were conducted twice a week for 20 minutes each over the course of 10 weeks. All of the students benefitted from feedback, regardless of form, as is evidenced by increases in the average CWPM, and decreases in IWPM. However, the results were interesting in that five of the six students made more gains in CWPM when they were given feedback on IWPM. The authors speculate that this is because students were provided with feedback on the words read incorrectly which might have been negatively reinforcing as opposed to being provided with how many words the student read correctly.

Chafouleas et al. (2004) also used an intervention consisting of a treatment package of RR alone, RR with Performance Feedback (RR/FB), and RR/FB with Contingent Reward (RR/FB/REW) to determine their effects on the ORF of three second and third grade students. Although all three phases induced gains in CWPM for all three participants, the RR condition induced the most robust effect for two of the three students.
Of importance though, these same two students made relatively few errors, suggesting that there would be no need to provide reinforcement or feedback for these students, simply practicing reading would suffice.

Telling students that their reading performance will be timed is another way of providing verbal feedback to students. Cates and Rhymer (2006) wanted to determine if explicitly timing students on their reading of Dolch sight words would have an effect on the student’s performance. The classroom teachers of four second- and third-grade general education students who were experiencing difficulty reading sight words provided the proposed intervention. The students read sight word phrases off of index cards while being timed. Initially, the students were unaware that they were being timed, but were told during the experimental condition that the students should read accurately as fast as they could and the teacher would be timing them. Sessions were conducted once a day for 12 days. The explicit verbal feedback that the student would be timed was shown to increase the number of Dolch word phrases read correctly during the experimental conditions. A reversal design showed that there were immediate increases in performance levels for all students when the experimental condition began, followed by immediate decreases when the students returned to baseline and did not receive any type of feedback. Thus, simply telling students that their performance will be timed served to increase the ORF rates of sight words.

*Providing error correction.* Martens et al. (2007) studied the effects of using a fluency-building intervention that consisted of three components: phrase drill error correction (PD), listening passage preview (LPP), and repeated readings (RR). In the PD portion, the students read a passage, and if they made a mistake, they had to repeat the
phrase that contained the word. The student simply followed along in the story as the researcher read it aloud in LPP, but the student and researcher both read the story two times each in the RR portion of the session. The second and third grade students improved greatly in ORF, measured in CWPM, but the progress of the second graders was more gradual. Specifically, the second grade students gained an average of 2.2 grade levels over the course of the intervention as opposed to a gain of 2.9 grade levels by the third graders.

Likewise, Begeny and colleagues (2006) found that RR and PD increased the ORF of a third-grade special education student. In fact, PD was equally and sometimes more effective than standard RR procedures. Specifically, the student’s data were more stable and he made less errors during the PD condition. The results in the RR and PD conditions were higher than in the reward condition, suggesting that students do not need rewards in order for RR interventions to be successful.

Another study that reported improvement in CWPM by second grade students also used error correction in addition to repeated readings (Nelson, Alber, & Gordy, 2004). The four students in this study progressed through a treatment package of error correction (EC), EC plus RR, and finally EC plus RR with previously read materials. The EC condition alone did not confer much benefit to the students as far as increasing the CWPM, but IWPM did decline in this condition. The decline in IWPM continued throughout the next two conditions in which RR was implemented and also served to increase the CWPM for all four students.

Promoting generalization. One of the most disregarded academic concepts is generalization. Generalization is the notion that a behavior is maintained over time, can
be seen in a variety of environments with different people, or even help bring about new behavioral responses. As the ability to generalize a skill is complex, it is something that needs to be planned for and taught (Alber-Morgan, 2010; Cooper, Heron, & Heward, 2007). Generalization of ORF skills is usually referred to as being able to read non-practiced passages in the clinical or classroom setting.

There has not been a lot of research conducted on generalization of reading skills, even though it is such an important topic. However, over the last decade or so, some researchers have begun to delve into this subject matter. Daly, Martens, Hamler, Dool, and Eckert (1999) used a brief multi-element design to determine the most effective ORF intervention for four students with reading difficulties. As part of this study, they used high content overlap (HCO) passages as one of the treatment components with RR or LPP. These HCO passages contained a considerable number of words from the instructional passages. There were various treatments in which the participants were involved, including: reward (RE), RR, RR with sequential modifications (RR/SM), LPP/RR, LPPP/RR/SM, and LPP/RR/EM (easier materials). The participants were provided one of the brief treatments on instructional level passages, assessed on the number of CWPM, and then assessed on a HCO passage. A notable difference between this and some other studies that assess generalization is that the participants were actually given instruction on the HCO passages if there were no increases in CWPM on the HCO passages after receiving the intervention on the instructional passages in the RR/SM condition. The researchers expected the participants to perform well on the HCO passages due to the overlap; however, they found that three of the four participants needed instruction on these generalization passages in order to increase the number of
Therefore, the researchers stated that one cannot assume that ORF behavior will be generalized.

As these and other researchers continued to investigate generalization of reading skills, Daly, Bonfiglio, Mattson, Persampieri, and Foreman-Yates (2005) wanted to refine what they had already learned about ORF generalization. This study included three boys in the third and fourth grade who had learning disabilities and were receiving special education services. Again, the authors used HCO passages, this time varying the difficulty level of the material (hard vs. easy) and whether or not the participants received instruction, as well as providing a tangible reward for meeting a predetermined criterion level of CWPM on HCO passages. The authors alternated the treatments and found differing patterns for the participants. For two of the participants, higher responding was produced in the easier condition with instruction, whereas the third participant had more pronounced increases in the harder materials with instruction. The key point is that generalization was obtained for all of the participants; however, it was only after instruction on the HCO passages. Therefore, educators may need to program common stimuli (Cooper et al., 2007) in order to achieve better results in generalization of ORF skills.

Ardoin and colleagues (2007) reported on the effects of using RR and PD to increase ORF of six third-grade students, this time also examining the effects of using multiple exemplars (ME) for promoting generalization. During the RR condition, students read the same passage a total of four times, whereas during PD, students read a passage once and then were provided with error correction on specific words and phrases within the passage. The ME condition was different in that students read passages A and
B twice, but the correction procedure was the same as in the RR and PD conditions. Each student was able to earn a token in all conditions by beating a previous number of CWPM and making less errors. Results revealed that the average gains in CWPM increased for every student across readings within the same condition (i.e., higher each additional reading). Five of the six participants’ results also showed evidence of generalization as the ORF rates on the generalization passages was higher than their first average of passage A (the passage with LPP). Similar generalization results were found in the ME condition, however, all students read less CWPM on the first reading of passage B as opposed to the second reading of passage A. This suggests that the LPP the students received before passage A was quite beneficial in boosting their initial fluency. Overall, both interventions were successful in increasing ORF, but three students clearly showed more improvement in the RR condition as opposed to the ME condition.

Ardoin et al. (2008) continued researching generalization, this time with 42 students in second and third grade. Again, the researchers used a RR condition with LPP and error correction. The procedures in the ME condition were identical to the RR condition, except that the participants read three different intervention passages once as opposed to reading the same intervention passage three times. Generalization was assessed immediately after each intervention condition (RR and ME). However, generalization passages in this study differed from the previous studies because the researchers used high and medium word overlap. Results revealed that the students’ ORF increased significantly in the RR condition. The authors hypothesize that this is due to the order of words and sentences staying constant during the RR condition as opposed to the order of words varying in the ME overlap passages. When examining ORF gains in
generalization passages, the RR condition was still effective, however, more gains in
generalization were noted in the medium word overlap ME condition. The authors
speculate possible reasons including that the medium word overlap passages actually
provided the students with opportunities to learn more words, thus increasing their ability
to read new passages more fluently. Overall, these studies reveal the effectiveness of RR
interventions to increase ORF and even generalization.

**ORF Interventions by Grade Level**

*Kindergarten/First.* Research on ORF tends to take place with students in the
intermediate or middle school grades, neglecting early elementary school children. This
might occur because the early years in school have always provided children with the
foundational skills for reading. Generally, preschool and Kindergarten teachers focus on
developing alphabetic principle and phonemic awareness skills of their students. They
continue to hone these skills in first grade. However, the little research that is available
for students in Kindergarten and first grade suggest that ORF instruction can begin earlier
than second grade (e.g., Hapstak & Tracey, 2007; Hitchcock et al., 2004; Lane et al.,
during the acquisition stage of reading, before reading fluency problems have a chance to
manifest themselves. Once these problems arise, they are likely to persist as children
progress through school (Speece et al., 2003), which is detrimental to the student’s
academic success.

One way to decrease the likelihood of future difficulty in ORF is by monitoring
the subword skills that have been demonstrated to be correlated with later reading
achievement as Speece and colleagues (2003) have done. In a descriptive longitudinal
study of fluency subskills that might predict early success in reading, these authors have demonstrated that letter-name fluency (LNF) and NWF in Kindergarten could be valid measures of predicting subsequent success in ORF in first grade. Moreover, NWF was the strongest predictor of reading achievement in Kindergarten. Likewise, Speece and Ritchey (2005) noted that letter-sound fluency (LSF) in Kindergarten is a good predictor of ORF in first grade.

With a solid foundation in the above-mentioned basic reading skills, students will be capable of reading words, phrases, and sentences more accurately; and therefore can begin fluency training. Vadasy and colleagues (2005) implemented an intervention with 57 first-graders who were at-risk for reading problems that consisted of three groups: Word Study (WS), Reading Practice (RP), and control. The WS and RP groups received explicit instruction in alphabetic principle and fluency, but the control group only received the normal classroom instruction. A comparison of pretest and posttest scores of students by group indicated that the RP and WS groups outperformed the control group on reading accuracy, comprehension, passage reading fluency, and spelling. One can infer, then, that more explicit instruction results in higher gains. The authors also indicated that the RP group made higher gains in reading accuracy and passage reading fluency than the WS group, which supports the idea that students indeed need to practice reading in connected text to increase their overall ORF.

Much in line with that thought, Hapstak and Tracey (2007) used an assisted-repeated reading intervention with leveled books for first grade students of varying reading ability. The repeated reading procedure involved (a) the student reading an initial passage, (b) teacher modeling of the reading while the student listened, (c) echo reading,
(d) two repeated readings by the student, and (e) final reading. The students were able to record the CWPM on a graph during the initial and final readings. The results indicated that this procedure was effective for increasing the CWPM for each student by session and for the entire study.

Second/Third. More research in ORF has taken place with the second and third grade population than with Kindergarten and first grade. The vast majority of this research has centered on the use of repeated readings (LeVasseur et al., 2008; Welsch, 2007), many of which include RR as part of a treatment package (e.g., Chafouleas et al., 2004; Martens et al., 2007; Nelson et al., 2004). Other reading research carried out with second and third grade students use methods such as peer tutoring (Barton-Arwood, Wehby, & Falk, 2005; Kourea, Cartledge, & Musti-Rao, 2007) and DI (Barton-Arwood et al.).

As noted previously, researchers have used RR to enhance the ORF of struggling readers. Martens et al. (2007) implemented a fluency-building intervention with second and third grade students and their results support the use of fluency instruction in the primary grades as most of the students responded well to treatment and made gains in ORF in both trained and untrained passages, although most students in the control group did not make gains on untrained passages. The participants in the Nelson et al. (2004) study had robust gains in CWPM. Two students gained an average of approximately 29 CWPM from baseline to the EC plus RR with previously read materials, and the overall group mean went from 50.33 to 72.15 CWPM, respectively. All participants in the Chafouleas et al. (2004) study also made gains over baseline compared to the three different phases. For the student that made the highest number of errors, the RR/FB and
RR/FB/REW phases brought about the most gains, with RR producing the least. The opposite was true for the two students that made the least number of errors, suggesting that the more errors students make the more feedback and reinforcement they require.

Response-time feedback was included as part of a RR intervention comparing three different approaches to increase the ORF and comprehension of second grade students (LeVasseur et al., 2008). The researchers were interested in how ORF would be impacted by the RR of word lists versus the RR of standard and cued text (passages presented to students as meaningful phrases or clauses, separated by extra spaces). In this way, students are prompted to read more prosodically, like they are speaking. During each condition, the students received feedback about their reading rate. The results were generally quite robust, as significant effects were obtained for many measures. They noted an increase in CWPM in standard and cued text as opposed to word lists, a decrease of IWPM across conditions, and an increase in ORF of cued over standard text and word lists. These results suggest that while reading word lists repeatedly can affect ORF in a positive manner, reading connected text repeatedly actually produces greater gains in ORF. Specifically, reading cued text can increase not only the accuracy and speed at which one reads, but also the ease of reading, or prosody.

Fluency research in the primary grades is scant. Research regarding fluency instruction in Kindergarten and first grade often focuses on the predictive ability of reading subskills in determining the subsequent reading level of students (e.g., Speece et al., 2003; Speece & Ritchey, 2005). However, researchers are starting to focus efforts on implementing actual fluency related interventions with this population (e.g., Coyne, Kame’enui, Simmons, & Harn, 2004; Hapstak & Tracey, 2007; Hitchcock et al., 2004;
Lane et al., 2007 Vadasy et al., 2005). Preliminary results seem promising (i.e., increases in CWPM and comprehension skills, and decreases in IWPM). The main interventions used with second and third grade students have been repeated reading and peer tutoring (e.g., Kourea et al., 2007; Lane et al.; Martens et al., 2007). Results obtained show promise as students in the all of the studies evidenced improvement in reading ability. Thus, fluency interventions should be continued with this age group and younger students as well.

Level of Reading Development and Disabilities

Much research has been conducted on students with reading problems. Some educators might believe that different interventions or instructional strategies might work well for one population as opposed to another. However, a review of the current literature on ORF suggests that the most promising and effective strategies work for students who are at-risk for developing a disability, students with mild-to-moderate disabilities, as well as students without disabilities, although results may vary. The NRP (2000) actually noted that among the studies they reviewed, differential effects of fluency interventions were obtained depending on the reading level of the participants. This means that students across ability levels benefit differentially from various types of ORF instruction, sometimes in idiosyncratic ways.

That point is evident in research by Hapstak and Tracey (2007). Although the assisted-repeated reading intervention was effective for all four first-grade participants, the data show support for use of this method particularly with those students who have difficulty with decoding. In this study, the student with a specific learning disability (SLD) and the at-risk student showed markedly more improvement in CWPM as opposed
to the ELL and general education student. Because the ELL had difficulty in oral language and the general education student read at an average to above average rate, the intervention was not as effective as it was for the other two students with difficulty decoding. That does not negate the fact that the assisted-repeated reading intervention still produced gains for the former students. Similarly, results from LeVasseur et al. (2008) indicated that RR training was valuable for all participants, but the most value was gained by the weakest readers. They noted that students with lower scores on the standardized reading fluency subtest actually made greater gains in CWPM and ORF, and had lower numbers of IWPM.

Martens et al. (2007) also saw favorable gains in the ORF of second and third grade students in their after-school intervention program. The study included students with and without disabilities. The authors stated that although the second and third grade students made gains in ORF, those who benefited the most were less fluent readers to begin with (i.e., reading less CWPM). Furthermore, the authors noted that fluency instruction might be more beneficial to students who read above the preprimer level or have a solid foundation in early reading skills because the second grade students took longer to reach the reading criterion than third graders. That viewpoint is consistent with those of Chafouleas et al. (2004) who noted that the two at-risk students who showed the strongest reading ability with fewer IWPM during baseline benefited the most from the RR condition alone. Those results support the notion that students who read accurately, but slowly, might benefit from merely practice.

Welsch (2007) wanted to determine effectiveness of using experimental analysis to ascertain what type of reading intervention would be most beneficial to increase the
ORF of four students with SLD. He used an adaptation of an alternating-treatment design in which there was a brief analysis, followed by an extended analysis, and finally a best treatment phase. Data showing increasing CWPM in the extended analysis condition for three of the students seemed to indicate that RR/EM was the best treatment, while RR produced greater average gains for the other student. The author noted increasing trends for instructional passages in the best treatment condition for all students, with corresponding decreasing or steady, low level trends in IWPM.

These findings suggest that the best intervention to teach ORF can be determined on an individual basis, thereby ensuring that all students receive empirically-validated interventions to increase their academic success (Welsch, 2007). In addition, various methods appear to be differentially effective for all students, regardless of the presence or absence of a disability, so if teachers implement one intervention and the corresponding data do not reveal effects for that individual, there are other evidence-based reading interventions that can, and should be implemented.

*English Language Learners (ELLs)*

Children who are not native English speakers often struggle with developing language and reading skills, including ORF. In the past, the educational topic of debate concerning ELLs has centered on the language of instruction. Now, however, more research is focused on what effective instruction looks like for ELLs (Denton et al., 2004; Gersten & Baker, 2000; Huebner, 2009; Linan-Thompson et al., 2003; Teale, 2009). Furthermore, ELLs are just as susceptible to being unfairly placed in special education as other minorities (Haager & Windmueller, 2001; Linan-Thompson et al., 2007). As such,
current researchers are examining ELLs’ response-to-intervention ([RtI], Haager & Windmueller; Linan-Thompson et al., 2007; McIntosh, Graves, & Gersten, 2007).

Linan-Thompson and colleagues (2007) point out that more research is needed on RtI with ELLs because this population has generally not performed well on academic achievement tests, and are also misidentified as needing special education services. However, the authors note that RtI is difficult to determine for ELLs because there is no standard or consensus on how to determine responders versus nonresponders. McIntosh and colleagues (2007) provide an account of how first-grade ELLs responded to tier one and two interventions implemented by classroom teachers. The researchers found that the students in the more effective teachers’ class made more gains in ORF during their second year of intervention, and these teachers provided more small group and individualized reading instruction. Furthermore, only nine students were labeled with a disability by the end of third grade. Eight of these students read only 20 CWPM or less at the end of first grade. Such a low reading rate is quite bothersome for any child at the end of first grade, but this is especially bothersome for ELLs who already have difficulty speaking English. For that reason, it is imperative to provide ELLs with appropriate classroom instruction and effective supplemental reading interventions.

Some recent researchers have studied the effects of commercial reading programs on ELLs’ reading skills. Santoro, Jitendra, Starosta, and Sacks (2006) used Read Well with four second-grade ELL students. Unlike other studies of mainly Spanish speaking ELLs (Denton et al., 2004; Haager & Windmueller, 2001; Linan-Thompson et al., 2003), this study included students of Middle Eastern, Hispanic, and Asian descent. The
intervention lasted 7-14 weeks. All participants increased their ORF from baseline to intervention, although differential effects were noted for the participants.

Another study (Denton et al., 2004) used Read Well and Read Naturally (ME) to determine their effectiveness on the reading skills of bilingual students. Participants included 93 students in second through fifth grade. Tutoring occurred three times a week for 40 minutes over 10 weeks. Results revealed that the Read Well program significantly increased the tutored students’ decoding, but there were no significant differences between groups on comprehension measures. The students in the RN program made negligible gains in comprehension and little to no gains in decoding. However, the authors speculate that the lack of comprehension effects might be due to an increased focus on ORF in the RN program, and for ELLs who are not reading in their native language, more time may be needed to understand what they have read.

Reading failure is an issue for native English speakers and ELLs alike. However, the above research demonstrates that the instructional techniques and interventions used for native English speakers are also useful for ELLs. Therefore, educators must use this knowledge to provide effective reading instruction for their whole student population.

Oral Reading Fluency: Implications for Practice

Teach ORF early. Given that fluency is complex and is comprised of every reading process and subskill (Wolf & Katzir-Cohen, 2001), it is imperative that students are monitored on their proficiency of prerequisite skills during early schooling. The importance of appropriately assessing reading related skills cannot be underscored enough as they have the power to predict future reading success or failure. Indeed,
predicting future reading achievement is essential so that the gaps between low- and high-achieving students can be closed.

The limited research that is available regarding teaching reading fluency for students in the early primary grades shows promise (Hitchcock et al., 2004; Lane et al., 2007). Samuels (1979) noted that educators can instruct students in word recognition at the accuracy level. It is during this accuracy stage when students are learning to read words but are devoting much attention to decoding, causing them to read more slowly. This stage seems to be an opportune time to begin fluency instruction. Moreover, teaching fluency skills early, before students actually fall behind their classmates, is quite logical considering it will save time, money, and effort (Coyne et al., 2004) plus provide students with early opportunities to achieve academic success. As Torgesen (2004) notes, prevention as opposed to remediation efforts, tend to decrease gaps between struggling readers and typically developing students because students receive instruction before too much time passes in which the students are not reading as much as their peers.

*Program for and measure comprehension.* An area that has not been widely researched in regard to ORF is comprehension. While many studies have included comprehension outcome measures, some noting positive effects others noting negligible effects, research needs to be conducted to determine the point at which ORF actually impacts comprehension. In any case, there is definitely a link between ORF and overall reading competence, or comprehension. According to Eckert et al. (2006) students must exhibit strong ORF skills in order to develop good comprehension; and increasing a student’s reading rate (fluency) will likely increase reading comprehension (Sorrell, Bell, & McCallum, 2007). However, one cannot assume that an increase in ORF automatically
promoted an increase in comprehension. Teachers must teach students how to comprehend reading material and provide many opportunities to assess reading comprehension.

*Program for prosody.* The importance of prosody as a component of fluent reading cannot be dismissed. Of the reviewed studies, only one actually included a measure of prosodic reading as part of the intervention (LeVasseur et al., 2008). Researchers have provided educators with a list of ways to assess prosody in addition to various methods that focus on increasing prosodic reading (Hudson et al., 2005; Rasinski, 2006). Therefore, it is incumbent upon educators to utilize such information to assist their students in becoming more fluent readers.

*Provide solid reading instruction for urban learners.* Evidence of reading failure in our nation’s urban schools is provided by the TUDA. Strong evidence-based reading practices must be used to teach urban learners in order to help these students reach their academic potential. Undeniably, children in urban schools could benefit from explicit instruction in ORF and its prerequisite and subskills. These include teaching phonemic awareness and the alphabetic principle directly, explicitly, and systematically (Musti-Rao & Cartledge, 2007b). Also, identifying at-risk learners early and providing them with many opportunities to respond (i.e., peer-tutoring) (Musti-Rao & Cartledge, 2007a) could go a long way in helping urban learners make sufficient progress in ORF, which will likely have a positive impact on their overall academic performance.

*Employ evidence-based instruction.* Repeated reading is an evidence-based practice that has been shown to be effective across grade level, disability category, and intervention agents. Trained paraprofessionals delivered a RR intervention with first
grade students who were at-risk for reading failure that helped the students surpass the
group that did not receive the RR intervention on measures of oral reading (Vadasy et al.,
2005). Hitchcock et al. (2004) used community partners as tutors to implement an
intervention that included RR and video-self modeling. Through this method, the first
grade students were able to not only increase their ORF and comprehension, but were
able to maintain and generalize it as well. LeVasseur et al. (2008) actually implemented
RR instruction on the computer, and these second grade students showed great
improvement in CWPM in addition to decreasing their word reading errors. All of these
findings taken together reveal the effectiveness of repeated reading

Conclusions

Fluency is a necessary and vital component of being a successful reader. As
discussed previously, fluency instruction needs to begin in grades as low as Kindergarten,
once students have a grasp of phonological skills and the alphabetic principle. Wolf and
Katzir-Cohen (2001) take it one step farther remarking, “...efforts to address fluency
must begin at the beginning of the reading process – that is, during acquisition – not after
reading is already acquired” (p.219). As a matter of fact, Fuchs et al. (2001) noted that
students in the primary grades actually show the greatest growth curve for ORF. If such is
the case, then it is essential that fluency-oriented instruction begin early in a child’s
school career.

Basically, we need to do what works; what has been scientifically reported and
validated. Even G. Reid Lyon reported this same sentiment over 10 years ago, wherein he
noted that children should be educated by methods that have been shown to be
scientifically successful, and to rid ourselves of assumptions that are not successfully
educating our children. Lyon (1998) wrote: “Good readers are phonemically aware, understand the alphabetic principle, apply these skills in a rapid and fluent manner, possess strong vocabularies and syntactical and grammatical skills, and relate reading to their own experiences. Difficulties in any of these areas can impede reading development” (p. 16).

Even though the present review was limited to ORF interventions with the primary grades, they serve to strengthen the literature on effective methods that increase reading fluency. Most notably, fluency related interventions can be implemented successfully with youngsters as early as Kindergarten. That should serve as a catalyst to increase fluency-based instruction in primary classrooms across the country. As stated by McCardle and Chhabra (2004), “When the research provides convergent evidence of what works, for whom, under what conditions, it is important that that information be put into practice” (p.476).

Potential Benefits of Technology

As noted at the beginning of this chapter, a second part of this literature review will focus on computer assisted instruction to teach reading. This review will specifically consider computer programs that have been effective with various types of learners in urban environments. The benefits of using technology in classrooms will also be discussed.

There is great potential for the appropriate use of computers and technology to impact teaching, learning, and even the opportunities for work (Hess & Leal, 2001). Technology could play a valuable role for teachers in many ways. Using technology within a classroom could save instructional and preparation time for teachers (Lee &
Vail, 2005). It usually does not take a lot of time to train the students on how to use the computer program, and can be a great supplement to traditional classroom reading instruction. Computers offer flexibility without an additional time burden (Mathes et al., 2001), as teachers in the Cassady and Smith (2004) study commented on the ease of learning and implementing the Integrated Learning System in their class. Computer-assisted instruction even allows teachers to get better at monitoring their students’ progress, which could boost academic performance (McCullough, 1995). Furthermore, computer programs provide teachers with a model of effective pedagogy, which, in theory, should increase student achievement (Chambers et al., 2008).

Students can also benefit from technology in a variety of ways. The time benefit is not just for teachers. For example, in an average of just 25 minutes, Lee and Vail (2005) taught sight words to students with developmental disabilities in Kindergarten and second grade. Lewandowski and colleagues (2006) used CAI in three 10-minute sessions to help improve word recognition in lists and reading fluency in connected text for third-grade students in an urban school. Perhaps part of the reason that students benefit from CAI is due to the novelty of working with computers, or the fact that most students in today’s society grew up using computers that can offer what teachers sometimes cannot.

Computer-assisted instruction can provide active engagement and interaction (Lonigan et al., 2003); immediate feedback (Hall et al., 2000; Lonigan et al.; Sorrell et al., 2007); reinforcement (Lonigan et al.; Macaruso & Walker, 2008); and modeling (Lonigan et al., 2003). Students are often able to work at their own pace (Lonigan et al; McCullough, 1995) with interesting and motivating activities (Macaruso et al., 2006), and are provided with many opportunities to repeatedly practice skills to build fluency.
(Macaruso et al.) and mastery in an environment that is not threatening or embarrassing (Diem & Katims, 2002; McCullough). Integrating multimedia content during classroom instruction can make the content easier for students to remember and learn because the content is presented visually and auditorily (Chambers et al., 2008). All of the aforementioned factors might help students feel more confident in their own intellectual abilities.

Learning by CAI might be especially beneficial for students with disabilities. For example, students with physical impairments might be able to access literacy activities by computer more easily, as some programs simply require the push of a button or computer key to turn a page, instead of turning a page in an actual book (Parette, Hourcade, Dinelli, & Boeckmann, 2009). Highlighting information (Parette et al.) might also be beneficial as it makes certain words, phrases, or content more salient for the learner. Students with learning disabilities also benefit from corrective feedback that computers provided during CAI on word recognition (McCullough, 1995).

Provide more intervention to more students. Computer-assisted instruction can be of great help to all students, particularly those with disabilities, who need more time and practice to learn how to read (Hall et al., 2000). Lee and Vail (2005) agree that computerized instruction could provide students with disabilities more direct instruction and repeated practice while the students are mainstreamed in the general education class. The authors offer that computers might be used as a teaching tool for learning and practicing new vocabulary or concepts in the inclusive classroom. This might not be possible without the assistance of technology as an intervention for students with disabilities in the general education class. However, instruction should not just be limited
to students with disabilities. Teachers could actually instruct small reading groups while other students work on reading skills on the computer. Indeed, Cassady and Smith (2005) state that instead of the other students who are not in the small reading group completing “seatwork” at that time, the students could be completing individualized lessons on the computer. In this way, more students are able to receive more intervention on much needed reading skills.

With the addition of computer laboratories in many schools, teachers are now able to take their whole class to the computer lab and use computers to supplement their traditional classroom curriculum. Not only can students play educational games on the internet, but many schools have bought educational software and installed it on their network. That allows more students to access CAI during a short amount of time. In fact, students in an urban high school said that teachers were actually able to spend more time with each student individually when computers were integrated into classroom instruction (Diem & Katims, 2002). In like fashion, the NRP (2000) affirms that computers allow students more opportunities to interact with the material for an extended time as opposed to traditional classroom instruction. This is especially important for schools as Lewandowski and colleagues (2006) point out that students require a lot of instruction in a period of minimal resources for instructional staff.

*Popular Computer Programs*

With the plethora of computer-assisted reading programs available to educators, it is important to evaluate every aspect of a proposed program. Bishop and Santoro (2006) provide a framework for educators to use to evaluate early reading software prior to its use with at-risk students. The framework includes multiple criteria for each of three
categories: interface design, content delivered, and instructional design. The authors provide questions that can be asked in each category to determine the extent to which the proposed computer program meets these standards. The hope is that evaluating software using this framework will enable teachers to make educated decisions when determining which program to buy and how to use it within their classrooms (Bishop & Santoro).

Indeed, Mitchell and Fox (2001) point out the importance of teachers, researchers, and software developers understanding the potential benefits and limitations of technology use with emergent readers. Below is a table of some popular computer programs that educators might want to evaluate as research indicates these programs appear promising in promoting the reading skills of young learners.
<table>
<thead>
<tr>
<th>Name</th>
<th>Grade Level</th>
<th>Targeted Literacy Skills</th>
<th>Feedback Provided</th>
<th>Modifiable</th>
<th>Key Feature(s)</th>
<th>Empirical Support</th>
<th>Pros</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alphie’s Alley/Reading Reels</td>
<td>1-5</td>
<td>PA, ORF, Comp, Writing</td>
<td>Yes</td>
<td>Yes</td>
<td>Can link to other media content</td>
<td>Limited</td>
<td>Tutors and students can engage with the program</td>
</tr>
<tr>
<td>Clicker 5</td>
<td>Pre-K-5</td>
<td>PA, Phonics, Comp, Print Concepts, Word Recognition</td>
<td>Yes</td>
<td>Yes</td>
<td>Spanish and French options</td>
<td>Limited</td>
<td>Can operate by switches for S with PD</td>
</tr>
<tr>
<td>DaisyQuest and Daisy’s Castle</td>
<td>Pre-K-2</td>
<td>Phonological Awareness</td>
<td>Yes</td>
<td>No</td>
<td>Adventure game with clues as rewards</td>
<td>Yes</td>
<td>Decreases response time as S move through skill levels</td>
</tr>
<tr>
<td>Headsprout Early Reading</td>
<td>K-2</td>
<td>Print Awareness &amp; all 5 NRP components</td>
<td>Yes</td>
<td>No</td>
<td>Web-based, entertaining animation</td>
<td>Yes</td>
<td>Money-back guarantee</td>
</tr>
<tr>
<td>Lexia Learning (PBR, SOS, &amp; Early Reading)</td>
<td>Pre-K-8</td>
<td>PA, Phonics &amp; WA</td>
<td>Yes</td>
<td>No</td>
<td>Spanish option</td>
<td>Yes</td>
<td>Branches back for review &amp; mastery of skills</td>
</tr>
<tr>
<td>Read Naturally (SE)</td>
<td>1-8</td>
<td>ORF, Vocabulary, Comprehension</td>
<td>Yes</td>
<td>Yes</td>
<td>Non-fiction passages</td>
<td>Limited</td>
<td>Provides fluent model for reading</td>
</tr>
<tr>
<td>WERP</td>
<td>K-2</td>
<td>ILS of reading, writing, and typing skills (including all 5 NRP components)</td>
<td>Yes</td>
<td>Yes</td>
<td>Includes activities for a whole year</td>
<td>Yes</td>
<td>Includes take-home activities</td>
</tr>
</tbody>
</table>

Note. PA = Phonemic Awareness; Comp = Comprehension; WERP = Waterford Early Reading Program; PBR = Phonics Based Reading; SOS = Strategies for Older Students; ILS = Integrated Learning System; WA = Word Attack; PD = physical disabilities; S = students; NRP = National Reading Panel; SE = Software Edition; ORF = Oral Reading Fluency.

Table 2.1 Popular Computer Programs
Computer-Assisted Instruction and Reading Skills

Researchers have studied the effects of CAI with varied populations and on a full range of reading skills, including those found essential by the NRP. For instance, CAI in reading has been shown to help develop phonological awareness (Lonigan et al., 2003; Mathes et al., 2001; Mitchell & Fox, 2001); phonics (Macaruso et al., 2006; Macaruso & Walker, 2008); sight words (Lee & Vail, 2005; Lewandowski et al., 2006); and a combination of reading skills (Cassady & Smith, 2004; Cassady & Smith, 2005; Clarfield & Stoner, 2005; Hecht & Close, 2002; Knezek & Christensen, 2007), for students with disabilities (Clarfield & Stoner; Lee & Vail; Lewandowski et al.) and at-risk students (Lonigan et al.; Mitchell & Fox), as well as in whole classrooms (Hecht & Close; Macaruso et al.; Mathes et al.).

CAI: Current Research

These studies on CAI in reading are for students in preschool through second grade, using different types of computer programs. The programs have been used with a wide variety of students, including those with disabilities.

Students with disabilities. Computer-assisted instruction has the ability to enhance learning for students with and without disabilities. However, Lee and Vail (2005) note that there is limited research on the effect of technology-based instruction for students with disabilities. Therefore, the researchers used a computer program, Word Wizard, to teach sight words to four students with developmental disabilities in Kindergarten through second-grade. The program included video clips to enhance knowledge, and target sight words that were specific to each student. There was a 5-second constant-time
delay procedure in the computer program, and the students’ responses were automatically recorded and saved by the computer. Students in this study also read storybooks and words on index cards as generalization measures. The results showed that all four students with disabilities were able to learn their targeted sight words through the use of CAI. Maintenance effects were also noted for two of the four participants. The beneficial effects of the program can also be noted as all of the students generalized their knowledge on paper-and-pencil worksheets, and reading words on index cards and in storybooks.

Kindergarten and first grade students with Attention-Deficit Hyperactivity Disorder (ADHD) were also shown to have improved reading skills after receiving CAI. Clarfield and Stoner (2005) used Headsprout as an intervention to determine its effects on the ORF and task engagement of three boys, as opposed to small group instruction directed by a teacher. The students were observed in their classroom and in the computer lab to determine the effects of CAI on their on-task behavior. Measures of ORF were also assessed once a week throughout the study. The students received CAI during nonacademic time, three times a week, for 20-30 minutes per session. After just 8 to 10 weeks of CAI on Headsprout, all 3 students read more correct words per minute (CWPM) and even showed decreases in off-task behavior. The gains made in CWPM were educationally significant, as the authors point out that the weekly gains made in CWPM exceeded the average expected growth rate for general education students for 2 of the 3 students, and the third student’s gains exceeded those of students in special education.
The immediate decrease in off-task behavior once implementing the Headsprout computer program also indicates its efficacy as a highly engaging, interactive program.

*Urban schools.* Without a doubt, students in urban areas can benefit from the additional instruction that computers can provide. That might be part of the reason that the majority of the empirical literature on CAI is conducted in urban schools. Third graders at an urban school in New York took part in a word recognition CAI study by Lewandowski and his colleagues (2006). The students were randomly assigned to a computer, tutor, or control group, and were trained on word lists over the course of three weeks. There was a training word list that the students practiced three times during a session; with the last being timed and scored for accuracy. The generalization word list did not include any words from the training list and were specifically developed to assess generalization of word recognition skills. Reading passages from first through fifth grade were also used during pre- and post-testing. The experimental procedure varied slightly by condition. During the computer condition, the students looked at the word on the screen while the computer read the word aloud; students heard the word through headphones, and said the word silently. The difference in the tutor condition was that the tutor read the word aloud from the list; and during the control condition, the student had to read the word from the list without any help from the experimenter.

Results of the study revealed that the tutor and computer groups significantly outperformed the control group on word recognition, but there were no generalization effects for any group. For the reading passages, the tutor and computer groups significantly outperformed the control group on fifth grade passages; and both groups
performed significantly better at post-test over pre-test on second through fifth grade passages. When the authors analyzed fluency gains (in words correct per minute) by group, across all grade levels, the tutor and computer groups performed well, with approximately 42.3 and 47.6 words respectively, while the control group only gained 14.7 words. Lewandowski and colleagues (2006) speculated that the tutor and computer groups performed better than the control group because each condition used research-based methods (i.e., direct instruction, error-free modeling and sight-sound associations, etc.). Also of importance is that practice reading lists increased fluency as shown by a decrease in list reading time. One of the hypothesized reasons for lack of generalization effects on the generalization list was that the students were practicing sight words, not blending. But, the authors did note that generalization of training lists could be seen as the reading passages included words from the training list, and the students’ fluency on those passages increased at post-test.

Macaruso and associates (2006) investigated the effects of two computer programs by Lexia Learning Systems (Phonics Based Reading [PBR] & Strategies for Older Students [SOS]) on the reading skills of at-risk first-grade students in five different urban elementary schools. The control group continued to receive the traditional classroom reading instruction on explicit phonics. The treatment group received CAI 2-3 times a week for 20-30 minutes each, over the course of six months. The students in the treatment group had to progress through the PBR activities before advancing to the SOS program, of which only 14 students actually worked on the SOS program. Pre- and post-test performance was assessed using the Gates-MacGinitie Reading Test, Level BR.
The authors conducted two sets of analyses on these scores; one for all students in both groups, the other for only Title 1 students in both groups. For the analysis of all students, results revealed significant gains for students in the CAI treatment and control groups. The analysis of only Title 1 students revealed similar results with both CAI treatment and control groups scoring significantly better at post-test. However, when pre-test scores were used as a covariate, the results showed that the Title 1 CAI treatment group scored significantly better than the Title 1 control group on the post-test. Also of importance is that the Title 1 CAI treatment group actually caught up to the non-Title 1 CAI treatment group by post-test. These results reveal the usefulness of the CAI intervention, as at-risk students who received this intervention outperformed their counterparts who did not receive CAI; and even made enough progress to catch up to their peers who were not at-risk. Macaruso and colleagues (2006) did note that the most gains in reading performance were made for Title 1 CAI treatment students who completed the most amount of units in the program. This speaks to the need for regular use of supplementary computer-assisted reading instruction.

Due to the favorable results achieved with the first graders in the above-mentioned study, Macaruso and Walker (2008) conducted a study with Kindergarten students using a different Lexia Learning Systems program, *Early Reading*, which focuses on early literacy skills. Six matched treatment and control groups from two different urban elementary schools participated in this study. “Matching” allowed the researchers to hold the teacher, classroom, and instruction constant, which should allow for greater experimental control and establish a functional relationship showing that the
differences between the treatment and control groups were actually due to the CAI intervention itself.

As with the aforementioned study, the CAI intervention was implemented 2-3 times per week for 15-20 minutes each over the course of six months. This time, two different subtests of the *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS) were used as a pre-test. Post-test measures included DIBELS and the *Gates-MacGinitie Reading Test, Level PR*. Again, Macaruso and Walker (2008) ran two sets of analyses; one for all students, the other for low performers only. For the all students analysis, post-test DIBELS scores were not significantly different by group. Comparison of the *Gates-MacGinitie Reading Test*, using DIBELS pre-test scores as a covariate, revealed that the CAI treatment group did in fact score significantly higher than the control group. Similar results were revealed when analyzing the results for low performers. In fact, after analyzing effect sizes for all versus low performing students, the authors concluded that the treatment group outperforming the control group was greater for the low performers than the all students analysis. Taking a closer look, two-thirds of the low performers in the treatment group scored above average on the Gates-MacGinitie normal curve equivalent (NCE) in comparison to only 1 of 12 in the control group. As a whole, these findings reveal the benefit of systematic CAI on the phonological awareness skills of Kindergarten students, in particular, low performing urban students.

*Grade levels.* Computer-assisted instruction can be implemented across all grade levels. The focus of this section is on reading programs that can be implemented with
students in Pre-Kindergarten through second grade, as these are the grades when students learn the foundational skills needed to read.

It might be difficult for some to believe that students as young as four-years-old are capable of learning reading readiness skills by computer, but they are. In fact, Lonigan and colleagues (2003) conducted a CAI phonological sensitivity program with Head Start children who were at-risk for reading difficulty. The children used the computer programs *DaisyQuest* and *Daisy’s Castle* for 15-20 minutes, 4 to 5 days a week. Even though the children worked independently, there were research assistants who would help the children as needed. Pre-tests measures for the CAI and control group were comparable except for letter sounds; the control group knew significantly more letter sounds than the CAI group. After just 8 weeks of intervention, the children in the CAI condition outperformed control group children on phonological sensitivity tasks that were explicitly taught by the computer program (e.g., rhyme sensitivity, phoneme segmentation, etc.). What is important to note is that skills that were not explicitly taught by the computer (i.e., expressive vocabulary), did not necessarily increase for CAI students. Therefore, one can infer that the computer program was responsible for increasing the phonological skills of students in the CAI group.

Peer-Assisted Learning Strategies for First-Grade Readers (1st-Grade PALS) is an intervention program for students who require reading remediation. Its effectiveness has been documented, but Mathes et al. (2001) wanted to determine if adding 8-10 hours of phonological awareness CAI would increase this skill for low-achieving (LA) first grade students. High- (HA) and average-achieving (AA) students were also included in this
study. There were three groups: (1) PALS, (2) PALS + CAI, and (3) contrast; with 12 students per group. The PALS group received PALS only; in addition to PALS, the PALS + CAI group received three 20-30 minute sessions on Daisy Quest and Daisy’s Castle over the course of 8 weeks; and the contrast group received the normal classroom instruction. Multiple pre- and post-test measures were taken (i.e., CTOPP, WRMT-R, TOWRE, etc.) which revealed a number of findings. One particularly surprising result is that even though the LA students in the PALS + CAI condition tended to outperform their counterparts in the contrast group, there were generally no significant differences between LA student performance in PALS versus PALS + CAI. However, this should not imply that the CAI did not have an effect on the performance of LA students. The LA students in the PALS + CAI condition actually had an initial increase in phonological awareness skills, even though this increase was not maintained over the course of the study. The authors did speculate that the lack of long-term CAI benefits to LA students might be due to this group’s pretest scores actually being lower than those of the PALS only and contrast groups. Of course, further research would need to determine if this is the case.

Contrary to the study above, first-graders in a high poverty urban school were able to benefit from technology that was infused in their curriculum. Chambers and colleagues (2008) implemented two different types of technology within Success for All first-grade classrooms at two different schools. The first, Reading Reels, was an embedded multimedia program that teachers actually implemented in their regular classroom instruction. The second type, computer-assisted tutoring using Alphie’s Alley,
was implemented on an individual basis, 20-minutes daily, by the Success for All tutor. The technology and control conditions had three groups each, for a total of six different groups. Multiple dependent measures were taken (i.e., Woodcock-Johnson III, Gray Oral Reading Test) and results were analyzed by all students, tutored students, and nontutored students. The technology group outperformed the control group when all students were taken into consideration. Analysis of tutored students only revealed that the technology group scored significantly higher than the control group on all of the dependent measures. The nontutored group did not fare as well, as the technology group only scored significantly higher on three of the five measures. The authors note that even though statistical significance was not reached on all outcomes, embedded multimedia and computer-assisted tutoring still show promise as this technology helps students remember concepts better by using both visual and auditory content.

Another promising example of technology being implemented into everyday classroom lessons is that of the Integrated Learning Systems (ILS). Cassady and Smith (2004) had the rare opportunity to study the effects of ILS on the phonological awareness of Kindergarten students in two different schools; one suburban (School B) and one rural (School A). Both schools used the same type of reading curricula that included phonics and whole language instruction, but School B (the no-ILS group) had technological difficulties which delayed the implementation of the ILS program. Therefore, the authors grasped the opportunity to test the differences between the curricula with and without an ILS. School A (the ILS group), used the *Waterford Early Reading Program* (WERP) throughout the whole school year. Students used the WERP for 20 minutes per day, every
day. The teacher aligned the activities each student would complete according to their individual classroom goals. Thus, the activities were individualized and were not completed in the same predetermined order for each student. School B teachers taught the traditional classroom curriculum without integrating any technology.

The researchers collected data on measures of phonological awareness and print concepts three times that year for all students in both schools. Results indicated that all students in both groups showed increased scores on print knowledge, which Cassady and Smith (2004) expected as these concepts are taught in Kindergarten, but not specifically taught by the WERP. However, there was a statistical significant difference by group on measures of phonological awareness in which the ILS group learned these skills faster than the no-ILS group in the fall, and even maintained this lead until the end of the school year. One caveat is that teachers in both schools were receiving ongoing professional development from university researchers on implementing technology into their curricula, therefore, the teachers in the ILS group felt empowered to infuse the WERP into everyday lessons, and even individualize the activities for each student.

The WERP was also used in two different sets of schools, one rural and one urban, to determine how CAI in phonemic awareness might impact emergent literacy skills of Kindergartners. Hecht and Close (2002) randomly selected children from six different public schools, two of which were not using the WERP, so those students were used as the control group. Multiple measures such as phonemic awareness, letter knowledge, invented spelling, vocabulary knowledge, and print concepts were assessed during pre and post tests. For six months, students in the treatment group used the WERP
for 15-minutes per day as part of their normal classroom instruction. Students in the control group did not receive any type of intervention.

Not surprisingly, results revealed that the WERP group made significantly greater gains in phonemic segmentation and blending, and the control children did not make any improvement in phonemic awareness over the six months. The authors also determined that student performance at pretest on various literacy skills (i.e., letter and vocabulary knowledge) was associated with gains in phonemic awareness. An additional finding is that “treatment resisters” actually received 10 hours less time on the WERP than students who made substantial growth. Due to this finding, Hecht and Close (2002) propose that students who have literacy-related deficits should receive more time using the computer-assisted instruction. In spite of the treatment resisters, the authors do assert that this study provides evidence that students can increase phonemic awareness knowledge regardless of how it is delivered (i.e., teacher versus computer).

*Implications*

Some important considerations must be made when determining the use of CAI in reading. One consideration deals with the attractiveness of the program to the students. Students are more likely to work independently if they like the learning tasks (Lonigan et al., 2003); and will stay on task and learn reading skills if they are motivated to do so (Macaruso et al., 2006) In general, students are often more motivated when they have some control over the program. The NRP (2000) suggests using hypertext for this purpose. Computer programs that allow self-pacing not only gives students more control over their learning, but also affords them the opportunity to practice skills until they are
mastered. For example, preschool children were allowed to repeat modules in *DaisyQuest* in order to master some phonological sensitivity skills like rhyming (Lonigan et al.), before moving on to other foundational reading skills.

One must also consider if the computer program is able to be modified to meet individual learner needs. *Alphie’s Alley* is a flexible program that allows tutors to choose whether or not they want to follow the pre-specified sequence or modify the program to suit the individual needs of particular students (Chambers et al., 2008). In fact, McCullough (1995) mentions that the best academic results can be achieved when the teacher is able to put information into the computer system and act as an interactive consultant to the program. The WERP actually allows teachers to do just that and customize the program to meet the needs of each student, so the student is not completing a predetermined sequence of activities (Cassady & Smith, 2004). *Clicker 5* also has the ability to be modified for diverse learners. The program is so flexible that teachers can actually develop individualized lessons that teach emergent literacy skills to a variety of learners (Parette et al., 2009). Students can also incorporate their own voice and pictures in this program, which might make it even more attractive to students; hopefully increasing their levels of engagement with the literacy activities.

*Read Naturally* (Software Edition) is another customizable program. Teachers can individualize reading goals and levels, and the program actually places the student in the appropriate level when instruction commences. The teacher can also modify the instructional sequence by: omitting certain activities; changing the number of times for read along; increasing the speed of the teacher model during read along; changing the
time allotted for retell; etc. These options definitely provide teachers with the tools needed to make an impact on their students’ reading achievement.

Another consideration that is important for teachers is whether or not CAI will free up some time. Researchers have noted that by using CAI with general education students, teachers could actually spend more time with students who need more individualized instruction (Macaruso & Walker, 2008). The WERP allows teachers to individualize the activities by student (Cassady & Smith, 2004; Cassady & Smith, 2005), thus teachers are able to not only attend to the individual needs of the students using the computer program, but also attend to the needs of the rest of the classroom population. As teachers carry so many instructional burdens, CAI in reading seems like a viable option as an intervention for students with disabilities and those struggling to read (Lewandowski et al., 2006). By using CAI, teachers will be able to reallocate their time while still providing students with disabilities more opportunities to practice in order to be able to acquire and maintain reading skills (Hall et al., 2000). However, the likelihood that any computer program will be effective depends a lot on how dedicated the teacher is to teaching his or her students to read (Cassady & Smith, 2005).

Additionally, it is important that students are familiar with how to use a computer and access the software (Mitchell & Fox, 2001) before teachers allow students to access CAI independently. Many students enjoy independently learning or practicing skills on the computer, as Lewandowski and colleagues (2006) point out that computers enable students to practice reading skills free from evaluation and judgment that humans bring. In addition, the authors state that computers are accurate in their ability to do so, but
teachers have the ability to provide individual attention and support to the student at the same time. Oftentimes, support might come in the form of simply monitoring students and being available to assist them with any computer related issues. Sometimes, even when students are working independently, the teacher might need to help the student, as Lafferty (2004) points out that Odyssey Reading does not necessarily reteach incorrect answers, and the program includes Spanish words that are not easily translated by students. Indeed, McCullough (1995) maintains the notion of “high tech, high touch,” meaning that it is still important to have a human there to guide and encourage the students while using the computer.

Also, administrators need to make sure that they provide teachers with professional development in the areas of instructional and assistive technology for successful implementation of computer programs. McCullough (1995) notes that teachers need support in order for CAI to be most effective for students. Administrators should make sure educators systematically integrate technology-based activities into their curriculum (Diem & Katims, 2002; Parette et al., 2009) throughout the course of the school day. For sure, one cannot expect that what a student learns on the computer in an isolated setting will necessarily transfer to the classroom. McCullough asserts that for skills learned on the computer to generalize, one must plan for it and integrate it within teaching. One way to do this is to infuse technology while teaching reading through the use of multimedia content (Chambers et al., 2008). The authors note that embedded multimedia do not take the place of the teacher, rather supplements the teacher’s lesson with animated illustrations of concepts. Accordingly, teachers need to be provided with
professional development on how to incorporate technology and content within a lesson (Diem & Katims), thus allowing students and teachers to benefit from computer assistance. Professional development of that sort should be a collaborative effort among universities, school districts, and businesses (Walker, 1997).

Strong leadership is also a prerequisite for successful integration of computer technology. Administrators must collaborate with teachers to ensure that the computer program fits within the existing curricula and all stakeholders’ beliefs about best practices in teaching. Furthermore, administrators have to develop a support plan for implementation beyond the technical support offered by the computer system (Cassady & Smith, 2005). However, cost of implementation should not be an issue, as the NRP (2000) points out that computers are now cost-effective in comparison to earlier computer technology from more than a decade ago. If administrators are concerned with cost, a way to lower the costs of technology is ensuring that teachers embed it within their lessons as opposed to buying commercial programs. Indeed, Chambers and colleagues (2008) assert that teachers can use multimedia content (such as Reading Reels) that is cost-effective and does not detract from the program’s success.

Nowhere is this cost issue truer than for urban schools. Urban districts must use scarce resources wisely so that technology can be available for teachers and students (Walker, 1997). Many students being educated in urban schools, including ELLs, are not reading on grade level, so additional instruction and practice via computers might prove beneficial. However, educators must use software that is systematic so as to produce better results in reading (Hall et al., 2000). Hecht & Close (2002) offer that students who
are at-risk for reading failure can be provided with phonemic awareness instruction effectively by using CAI, so it is a cost effective way to teach the students, as opposed to paying more teachers or aides. In this way, urban schools might be able to provide more effective instruction to the students who need it the most.

Directions for Future Research

There is tremendous potential for CAI to have a positive impact on the academic skills of all students. Even though researchers have made great strides in developing evidence-based computer programs, there is still a wealth of knowledge that is missing in current empirical literature; an enormous amount of research that is yet to be conducted on CAI in reading. Specifically, are Integrated Learning Systems more effective than stand-alone computer programs that are not connected to classroom content but still teach the skills students need to master? How might ongoing professional development with classroom observations affect the efficacy of the program (Diem and Katims, 2002)? Also, how much time and effort is required of the teacher? Does the computer free up time for other tasks? Lee and Vail (2005) believe that future research is needed to determine the impact of teacher participation on the effectiveness of CAI. Is there a novelty effect that exists when students work with computers? Do initial student gains actually decrease over time (MacArthur, Ferretti, Okolo, & Cavalier, 2001; Mitchell & Fox, 2001), or can treatment resisters improve if they are provided with more time (Hecht & Close, 2002)?

More research also needs to be conducted with a wide range of age groups. Can results be achieved with pre-K and older students (Hecht & Close, 2002) who have
deficits in foundational reading skills? The NRP (2000) also propose that work needs to be done on how teachers might integrate Internet applications and other multimedia (i.e., hypertext) into their classroom instruction; mainstream populations, not just special education; and speech recognition. Certainly, the next generation of CAI in reading should involve speech recognition, as MacArthur and colleagues (2001) suggest that researchers develop programs that can recognize the student’s speech.

Conclusions

It is important to identify students at-risk for reading difficulty early, during Kindergarten, so prevention efforts can begin (Torgesen, 1998). Technology could play an important role in preventing reading problems, thereby reducing the need for remediation in later school years. However, there are no “easy fixes” that exist to remediate the reading performance of students who are at-risk for reading failure (Mathes et al., 2001). Accordingly, technology should not be seen as a panacea for students experiencing continual academic difficulty (Diem & Katims, 2002). Indeed, MacArthur and colleagues (2001) express “…cautious optimism about technology’s potential to improve the literacy skills and instruction of students with disabilities” (p. 298).

Computer technology has, however, been shown to successfully provide many types of reading instruction (NRP, 2000). One way to increase the efficacy of CAI is to ensure that programs are developed that accommodate student needs (Lee & Vail, 2005). The authors suggest that variables such as the number and/or difficulty level of questions that students need to answer, a variety of error correction procedures and consequences, speed demands, active student responding and attentional cues, and sound be built into
computer programs. The NRP notes that software should be adapted so that it can recognize idiosyncrasies in student responses and react appropriately. Also, when developing CAI programs, Larsen (1995) suggests considering how much control the learner has over the program, types of feedback from the computer, and how graphics and screens are used and designed. From a research perspective, McCullough (1995) cites numerous variables that will likely influence the level of success of computer-assisted reading instruction, including the amount of time the student is on task, the severity of the student’s disability, the student’s age, and the type of feedback, to name a few.

We need to face the fact that technology has played an important part in most people’s lives since the inception of the World Wide Web. Children today tend to be handier with technology than adults, and tend to use technology on a regular basis at home and school. It would be beneficial, then, to engage students in learning activities in a modality that most already find interesting. Computer-assisted instruction is an avenue that keeps students engaged in learning tasks, and as Belfiore et al. (2005) assert, when students make academics a priority, they choose to participate in such tasks. As a matter of fact, Diem and Katims (2002) assert that students should share what they know about technology with their teachers and become part of the educational technology experience. Technology is more complex than ever before and is an integral part of workplaces and society; therefore, it should become an integral part of every classroom in the country. Chambers and colleagues (2008) mention that, “For more than half a century, a technology revolution in education has been just around the corner, yet it never quite arrives” (p.1). Are we there yet?
Chapter 3: Method

Chapter three describes the methods and procedures of the study. It contains information on the participants and setting, materials, definition of the dependent and independent variables, general procedures, inter-observer agreement, procedural integrity, and social validity.

Participants

The participants in this study included six second-grade students that attended a charter elementary school in a large urban Midwestern area. All six students evidenced a need for an ORF intervention. The students all came from low socioeconomic homes, and one student was an English-Language Learner. One student had a speech impairment. The six students, five girls and one boy, were all in the same classroom. However, three students went to the first-grade classroom for reading instruction.
Selection Criteria

To be selected as participants in this study, students were first recommended by their teacher as students who needed additional instruction in reading fluency. The students in this school were tested three times a year on the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002). In addition to teacher recommendation, the students had to score in the at-risk or some risk category to be included in this study. They also had to test below the 30th percentile on the Word Attack (WA), Letter Word Identification (LWID), and ORF reading subtests on the Woodcock-Johnson Test of Achievement – Third Edition (WJ-III ACH; Woodcock, McGrew, & Mather, 2001). Parental permission was a final requirement for participation.

North River Academy

Tabitha. Tabitha was a 7-year-old African-American female in the second grade. She scored 17 CWPM (at-risk) on the DIBELS ORF sub-test on the winter benchmark assessment. She was placed in a Direct Instruction program (Reading Mastery Plus Level I) for the whole school year. Her teacher taught reading mastery plus for 30-35 minutes daily from September through December. From January until the end of the school year, she received reading mastery instruction Monday through Thursday. There were 12 students in her group.

Diana. Diana was a 7-year-old African-American female in the second grade. She scored 18 CWPM (at-risk) on the DIBELS ORF sub-test on the winter benchmark assessment. She was placed in a Direct Instruction program (Reading Mastery Plus Level I) for the whole school year. Her teacher taught reading mastery plus for 30-35 minutes
daily from September through December. From January until the end of the school year, she received reading mastery instruction Monday through Thursday. There were 12 students in her group. Diana is going to be retained and completing second grade again next school year.

*Javon.* Javon was an 8-year-old African-American male in the second grade. He scored 18 CWPM (at-risk) on the DIBELS ORF sub-test on the winter benchmark assessment. He was placed in a Direct Instruction program (Reading Mastery Classic, Fast Cycle Level) from the beginning of the school year until December. From December until the end of the school year, he was in Reading Mastery Level II. His teacher taught reading mastery for 30 minutes everyday of the week until January, and everyday except Fridays from January until the end of the school year. There were 10 students in his reading group. Javon and Quantasia are fraternal twins.

*Angela.* Angela was a 7-year-old second-grade Asian female who was an English-Language Learner (ELL) and receives speech services. She scored 30 CWPM (at-risk) on the DIBELS ORF sub-test on the winter benchmark assessment. She was placed in a Direct Instruction program (Reading Mastery Plus Level I) for the whole school year. Her teacher taught reading mastery plus for 30-35 minutes daily from September through December. From January until the end of the school year, she received reading mastery instruction Monday through Thursday. There were four students in her group.

*Quantasia.* Quantasia was an 8-year-old African-American female in the second grade. She scored 25 CWPM (at-risk) on the DIBELS ORF sub-test on the winter benchmark assessment. Quantasia was placed in a Direct Instruction program (Reading
Mastery Classic, Fast Cycle Level) from the beginning of the school year until December. From December until the end of the school year, she was in Reading Mastery Level II. Her teacher taught reading mastery for 30 minutes everyday of the week until January, and everyday except Fridays from January until the end of the school year. There were 10 students in her reading group.

Tamika. Tamika was a 7-year-old African-American female in the second grade. She scored 43 CWPM (at-risk) on the DIBELS ORF sub-test on the winter benchmark assessment. Tamika was placed in a Direct Instruction program (Reading Mastery Classic, Fast Cycle Level) from the beginning of the school year until December. From December until the end of the school year, she was in Reading Mastery Level II. Her teacher taught reading mastery for 30 minutes everyday of the week until January, and everyday except Fridays from January until the end of the school year. There were 10 students in her reading group. Tamika also receives speech services.

Teachers

NRA used Direct Instruction (DI) as their reading curriculum. DI is an explicit phonics program that focuses on decoding and comprehension. Teachers using DI give instruction in small groups based on ability. Therefore, based on the participants’ abilities and learning needs, the participants were assigned to small groups in two different classrooms for reading instruction. The second-grade teacher, Teacher 2, taught reading to Javon, Quantasia, and Tamika. The first-grade teacher, Teacher 1, taught reading to Tabitha, Diana, and Angela. Tabitha and Diana were in one reading group, and Angela was in a different reading group.
Teacher 1 and Teacher 2 were licensed teachers in early childhood P-3. Teacher 1 was an African-American female who had two years of teaching experience, both at NRA. She was the first grade teacher and the reading teacher for three of the participants in the study. Teacher 2 was a Caucasian female who had four years of teaching experience, with three years at NRA. She was the second grade teacher and taught reading to the remaining three participants. All of the participants received content area academic instruction from Teacher 2.
<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Race</th>
<th>Class</th>
<th>Grade</th>
<th>DIBELS ORF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabitha</td>
<td>7-9</td>
<td>AA</td>
<td>Teacher 1</td>
<td>2\textsuperscript{nd}</td>
<td>At-risk 17</td>
</tr>
<tr>
<td>Diana</td>
<td>7-4</td>
<td>AA</td>
<td>Teacher 1</td>
<td>2\textsuperscript{nd}</td>
<td>At-risk 18</td>
</tr>
<tr>
<td>Javon</td>
<td>8-6</td>
<td>AA</td>
<td>Teacher 2</td>
<td>2\textsuperscript{nd}</td>
<td>At-risk 18</td>
</tr>
<tr>
<td>Angela</td>
<td>7-9</td>
<td>Asian</td>
<td>Teacher 1</td>
<td>2\textsuperscript{nd}</td>
<td>At-risk 30</td>
</tr>
<tr>
<td>Quantasia</td>
<td>8-6</td>
<td>AA</td>
<td>Teacher 2</td>
<td>2\textsuperscript{nd}</td>
<td>At-risk 25</td>
</tr>
<tr>
<td>Tamika</td>
<td>7-7</td>
<td>AA</td>
<td>Teacher 2</td>
<td>2\textsuperscript{nd}</td>
<td>At-risk 43</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Age: refers to the age of each participant in years and months at the time of pre-testing (2/01/2010)
\textsuperscript{b}Race: refers to the racial background of the participants; AA= African American
\textsuperscript{c}Class: refers to the class that each participant attended for reading group
\textsuperscript{d}ORF: oral reading fluency classification and score

Table 3.1 Participant demographic information and DIBELS winter benchmark scores
Experimenter and Secondary Observer

The primary experimenter for this study was a doctoral candidate in Special Education in the School of Physical Activity and Educational Services at The Ohio State University. She earned her Bachelor of Arts in Psychology from The Ohio State University in 2002 and a Master of Education in Special Education from The University of Toledo (Toledo, OH) in 2005. The experimenter worked as a general education teacher for two years, and a special education teacher for three years, in a charter school in Toledo, OH. During her teaching career, she was trained in and implemented evidence-based academic instruction for students with mild-to-moderate disabilities. In 2007, the experimenter began her doctoral studies in Special Education at The Ohio State University.

The secondary observer assessed inter-observer agreement and procedural integrity. She was a first year doctoral student in the Special Education program at The Ohio State University. She completed her Bachelor of Science in Special Education in 2004 from OSU. She completed her Master of Education in Special Education in 2005 from OSU. After she completed her M.Ed., she became a special education teacher for students with emotional and behavioral disorders in the city of Columbus, OH for four years. During her teaching career, she implemented various evidence-based academic and social skill programs for students with or at risk for disabilities. The primary experimenter trained her as the secondary observer. She was trained on the criteria for
scoring words on treatment and generalization probes as correct or incorrect, and on the inclusion criteria for word retell.

Setting

North River Academy

North River Academy (NRA) was the setting for this study. North River Academy was a charter school located in a low SES neighborhood inhabited by mostly African-Americans. Ninety-nine percent of the school population received free or reduced lunch. There were approximately 165 students enrolled in Kindergarten through 6th grade. The student population was split evenly, with 50% of boys and girls. The total student population with a disability was 13%. The ethnic student background was: 85% African-American, 7% Multi-racial, 7% Caucasian, and 1% Asian-American. NRA’s eight teachers used DI as the core reading curriculum across all grades.

The study took place in the participants’ urban elementary school. Testing sessions were in a multi-purpose room that was outside of the classroom. The room contained a conference table, four chairs, a refrigerator, two desks, spools of paper, multiple televisions on carts, a lap-top computer cart, books and other educational material, and the materials needed for testing. Intervention sessions took place in the latchkey room next door to the participants’ classroom. The room contained four small tables and chairs, a refrigerator, two desks with chairs, a file cabinet, and various toys and games. There were few distractions, although teachers or students retrieved items from the refrigerator on a couple occasions.
Definition and Measurement of Dependent Variables

This research study had two types of dependent variables. The first dependent variable directly measured the relationship between the dependent and independent variables. The data for this type of dependent variable came from the Read Naturally program, AIMSweb progress monitoring stories, and from the classroom curriculum (DI stories). The RN treatment probes were from the RN program, and the AIMSweb and DI stories were used as generalization probes. The second type of dependent variable indirectly measured the relationship between the dependent variables and independent variable. The pre- and post-test scores on the WJ-III ACH sub-tests (i.e., ORF, passage comprehension, letter-word identification, and word attack) and the pre- and post-test scores on the DIBELS ORF sub-test were used as data for the secondary dependent variables.

Primary Dependent Variables

Correct words per minute. Words were marked as correct if the student read the word correctly within 3 seconds, or self-corrected within 3 seconds. Dialect was also taken into account. Words were scored as correct if the student said the correct word but pronounced it slightly different, for example, using African American English ([AAE], Leslie & Caldwell, 2006). These data were collected during one-minute readings of connected text in the RN program and generalization passages.

Incorrect words per minute. Words were scored as incorrect if the student did not read the word within 3 seconds of the last word, or if they mispronounced or omitted a word. Additions and substitutions were also marked incorrect. These data were collected
during one-minute readings of connected text in the RN program and generalization passages.

*Comprehension questions.* The participants answered five questions about each story; four multiple choice questions each having four choices, and one short answer question. Question 1 asked about the main idea of the story; question 2 asked the participant to recall a literal fact; question 3 referred to a vocabulary word; question 4 asked the reader to make text connections; and question 5 was an open-ended, sometimes inferential question, about the author’s ideas. These data were collected after the participant reached his or her goal criterion for each specific story, and after every generalization probe.

*Oral retell.* The participants told the experimenter everything that he or she could remember from the story that was just read. Words were counted as correct if: a) it was in or directly related to the passage; b) it was part of a whole statement; and c) was not repeated. Non-words such as “um” or “hmm” were not counted. These data were collected after one-minute readings of connected text in the RN program and generalization passages.

*Secondary Dependent Variables*

*DIBELS.* North River Academy measures student performance on DIBELS benchmark assessments three times a year. Teacher 2 provided the experimenter with the DIBELS ORF data for each participant. This sub-test measures a child’s ability to read connected text. The participants were timed for one-minute and the teacher calculated the number of CWPM and IWPM and reported the median score to the experimenter. The
experimenter also gathered these data at the conclusion of the study as part of the post-test.

*WJ-III ACH.* The reading subtests of the WJ-III ACH were administered by the experimenter as a pre-/post-test measurement procedure to note any gains in reading ability. The subtests of the standard battery of the WJ-III ACH that were administered included letter-word identification (LWID), word attack (WA), passage comprehension (PC), and reading fluency (ORF). To determine the most accurate information about the reading level of the participants, the raw scores and age- and grade-equivalent scores were used. These scores then helped determine each participant’s standard scores. Form A was used for pre-testing and Form B was used for post-testing to minimize practice effects for the participants. The experimenter used the ceiling to score the LWID, WA, and PC sub-tests, which means that testing ended when the participant answered the six highest test items incorrectly. The experimenter used the 3-minute time limit to score the ORF sub-test.

**Definition and Measurement of Independent Variable**

The independent variable in this study was a computerized reading program, Read Naturally. This program had a pre-specified sequence of activities that helped build oral reading fluency. The participants followed the sequence that included: key words, a cold timing, read along, practice reading, quiz questions, retell, and pass timing. A detailed explanation of each variable follows.
Each experimental session followed the computer sequence of the RN program, including the activities stated below. Sessions were individual or small group, depending upon the classroom schedule and attendance of the participants.

*Key words.* The computer presented vocabulary words from the story by reading the word and then providing a definition of the word. Oftentimes, the program would also provide a picture and a sentence using the word. The program instructed the participant to read along silently and click on any word that the participant needed to hear repeated.

*Cold timing.* In this activity, the computer instructed the participant to read the passage for one minute and click on any words that the participant did not know. The words that the participant clicked became maroon in color and the computer counted them as errors. At the end of the one minute, the participant had to click on the last word read. After that, the computer instructed the participant to click on the words that he/she did not know to have the words read to them. The computer then calculated the CWPM and errors, and determined a goal for that story for the participant. However, for the purpose of this study, the goals for each participant were individually determined. For example, the first goal during the first criterion was based on the average of the two highest baseline readings. Goals for subsequent criterion changes were based on the two highest treatment probes from the previous criterion.

*Read along.* The participant was instructed to read along silently with the teacher model reading on the computer. The program highlighted sentences as they were being read, and the participant followed along. The participant read along with the computer three consecutive times. At the end of each reading, the vocabulary words were in blue
and the computer instructed the participants to click on the words they do not know to hear what they mean. The computer then pronounced the word, read the definition to the participant, and sometimes provided an example sentence using the word. After the participant read along with the computer three times, the program presented the option to click next to move on to practice reading, or read along with the computer again.

**Practice reading.** This activity allowed the participants to time themselves reading the story in order to meet the predetermined goal. The computer instructed the participant to click on a clock at the top of the screen which started the one-minute timing. The participant was allowed to click on a word he/she did not know or was having difficulty with, and the computer would read the word to the participant and score it as incorrect. At the end of one minute, a bell sounded and the computer advised the participant to click on the last word read. The computer then calculated the CWPM for the participant and displayed this number in a box on the right side of the screen. If the participant did not reach the goal, he/she read the story again. The participants practiced reading as many times as necessary in order to meet or exceed their goal. After the participant reached his/her goal, the computer instructed them to either click next to answer the comprehension questions, or continue practicing the story. According to RN, students usually practiced reading between three to ten times per story. In the current study, students practiced as few as one time, or as many as 18 times, to meet or exceed their goal. The length of these intervention sessions was between 20-45 minutes.

**Quiz questions.** After the participant reached the goal, a quiz appeared on the screen. The quiz consisted of five questions; four multiple choice and one short answer.
The participant clicked on the desired response for the multiple choice question, and dictated to the experimenter the answer for the short answer question. The experimenter typed it into the box on the computer because the second-grade participants did not have adequate typing skills. The program presented again any responses that the participant did not answer correctly; however, only the first set of responses was recorded on the data sheet.

Retell. The next activity allowed the participant to retell as much of the story as he/she could. The retell activity in the RN program allowed students to type as much as they can remember about the story. However, due to the limited typing ability of second-grade students, the experimenter asked that the participants completed an oral retell of the story. This means that the experimenter asked the students to orally state as much of the story as they could remember, while the experimenter used a counter to count the number of words retold in one-minute. The experimenter then wrote the final number on the data collection sheet.

Pass timing. The last activity in the sequence was pass timing, in which the experimenter entered a password in order for the participant to access the story. This activity was identical to the cold timing, and the scores during this pass timing were used for treatment probe data.

General Procedures

Experimental Design and Conditions

A multiple baseline across participants with changing criterion tactics design was used for this study. A multiple baseline across participants design is used when two or
more subjects are targeted for the same behavior (Cooper et al., 2007). This design consisted of three tiers with two participants in each tier for a total of six participants. All three tiers began in baseline and entered into treatment phases as the study progressed. Each participant entered into a new phase of the intervention when a predetermined criterion of CWPM had been reached. Visual inspection of the data was used to determine when a tier was introduced to the treatment and when the individual criteria needed to be increased.

The changing criterion design is used to determine the effects of a graduated or stepwise intervention on a single target behavior (Cooper et al., 2007). According to Cooper and colleagues, in order to properly implement a changing criterion design, three factors must be carefully manipulated: phase length, magnitude of changes, and number of criterion changes. However, the authors note that there are few published examples of pure changing criterion designs; instead, some researchers have used a changing criterion tactic to analyze data within a larger design. Cooper and colleagues recommend that the data should guide decision making as far as the length of phases, but the current investigation cannot be considered pure because the experimenter had a predetermined number of sessions in which the criterion had to be reached before changing to the next phase. Furthermore, the authors suggest that the size of the criterion changes should be varied. The experimenter in this study could not vary the magnitude of these criterion changes in advance, as these changes were based upon each participant’s individual performance.
Using this changing criterion tactic within the larger multiple baseline design allowed the experimenter to analyze the effects of increasing CWPM on generalization. This design was chosen because learning to read is irreversible and it did not require the intervention to be withdrawn. Additionally, embedding changing criterion tactics allowed for a series of treatment phases in which the participants had to show improved performance in order to advance to the next phase.

Each participant had individualized criterion for the CWPM before changing to the next treatment phase (criterion). The participant had to read at or above the individualized criterion for at least two sessions (data points) before advancing to the next phase. The criterion for phase one for each participant was based on the average of the top two baseline probes. The criterion for each additional phase for each participant was based on the average of the top two treatment probes from the previous phase. Systematically increasing the criterion allowed the experimenter to observe the effects of increased ORF on comprehension and generalization measures.

Pretest

Information from each participant’s DIBELS ORF test was gathered from the school’s pre-existing data. Students had to score in the at-risk or some risk category to be included in this study. The experimenter also gathered these data at the conclusion of the study as part of the post-test. Additionally, the reading subtests of the WJ-III ACH were administered by the experimenter as a pre-/post-test measurement procedure. The results of the pre-test were compared to the post-test results to note any gains in reading ability. The subtests of the standard battery of the WJ-III ACH that were administered included
letter-word identification (LWID), word attack (WA), passage comprehension (PC), and reading fluency (ORF). In order to determine the most accurate information about the reading level of the participants, the raw scores and age- and grade-equivalent scores were used. Pre-testing took place approximately one week prior to the implementation of baseline.

**Baseline**

Baseline for all students began approximately a week after administering and scoring the pre-tests. In baseline, each participant was presented with a RN passage that had been re-typed in a Word document on the computer. The participant read for one-minute to determine the number of CWPM and IWPM. The participant was given the passage and instructed to read as many words as he/she could in one minute. The experimenter stated, “This is a story that I would like you to read. Please try to read as many words as you can as quickly as you can until the timer goes off. Once you hear the timer, stop reading. If you do not know a word, I will tell it to you so that you can keep reading. Try to remember as much as you can about what you read because I will ask you to tell me about the story when you get finished.” The experimenter then set the timer to countdown one minute and started the timer once the participant read the first word. If the participant did not read the first word within 3 seconds of telling them to begin reading, the experimenter gave them the word and told them to continue reading. That word was marked incorrect. After the timer sounded, the experimenter recorded the number of CWPM and IWPM on the data collection sheet. The participant was then asked to retell as much of the story as could be remembered. A corresponding Word document with
comprehension questions was then be presented to the student, on the computer, to measure their level of comprehension on the passage they just read. The experimenter stated, “Now you are going to read these questions about the story. Tell me what you think the best answer is to each question and I will write it down for you.” The answers were recorded on a separate sheet of paper and the number of questions answered correctly was recorded on the data sheet. Baseline probes continued in this manner until a steady state of responding was achieved. A minimum of five baseline data points was required.

Training

Immediately after baseline, the participants needed to be trained to use the RN program. The experimenter trained each participant one-on-one on the entire sequence of one story. During the training, the experimenter taught each participant how to follow the story sequence that included key words, a cold timing, read along, practice reading, quiz questions, retell, and a pass timing. The participants were allowed to ask questions at any time, and were considered trained if he/she could go through an entire story sequence without prompting from the experimenter. Five of the six students only required one training session. Angela, the English-Language Learner, required two training sessions. A procedural integrity training checklist was also used in training (see Appendix E) to ensure that the experimenter properly trained the students on how to use the RN program. A secondary observer scored the training checklist.
Implementation of the Independent Variable – Read Naturally

After a stable trend in baseline and training on the RN program, Tier 1 entered into intervention. Every intervention session followed the same procedure. The participants sat in front of their respective computers and were told to enter his/her password into the RN program. The participants were then told to select a particular story and follow the instructional sequence. The participants engaged with the program independently, but were allowed to ask questions if needed. The experimenter only interacted with the program when she had to type the participant’s answer to the fifth comprehension question on the quiz; input a password before the pass timing (treatment probe); input the number of IWPM on the treatment probe; and input a score for prosody. Intervention sessions continued in like manner throughout the course of the study.

Each participant had a specific, individualized goal of CWPM that needed to be reached in order to change criterions. The first criterion for each participant was determined by calculating the average of the top two baseline probes. Each additional criterion change was calculated by averaging the two highest probes from the previous criterion. All data were visually inspected and discussed during weekly meetings with the experimenter, her advisor, and another doctoral student. Decisions about when to change each tier’s respective criteria were made after visually inspecting the graphs.

Treatment Probes

The last activity of the RN sequence, pass-timing, was used as the treatment probe. The participant was expected to complete a one-minute reading of the passage, and an oral retell of the story. The experimenter instructed the participant to read the
story on the screen and gave the same instructions provided during baseline. The participant clicked the start button on the screen and read the story for one minute. The computer then instructed the participant to click on the last word he/she read when the bell sounded. The experimenter then instructed the participant to retell as much of the story as he/she could remember. The experimenter started the timer for one minute and instructed the participant to stop retelling the story when the timer went off. After that, the experimenter input the number of errors in the RN program, and the computer then automatically calculated the CWPM that the student read during the pass timing. The experimenter then documented the number of CWPM, IWPM, comprehension questions answered correctly, and number of words retold on the data collection sheet.

Generalization Probes

After two treatment probes, each participant was given a generalization probe. Collection of generalization probes followed the same procedure as treatment probes. The first measure of generalization was non-trained passages given in the clinical setting (i.e., the latchkey room). For every two treatment probes, the participants read a new story from AIMSweb and the experimenter took data on all of the dependent variables. The second generalization measure was a probe from the classroom reading materials, Reading Mastery (Engelmann & Bruner, 1997). The participants were given these untrained passages in their regular classroom setting. However, the experimenter only probed the classroom materials for generalization five times throughout the course of the study. Comprehension questions for both generalization measures were written by the experimenter based upon the same format as the RN program.
Posttest

The study lasted approximately 16 weeks. Tiers 1, 2, and 3 received 12, 10, and 7 weeks of intervention, respectively. The same four subtests (WA, LWID, PC, ORF) of the WJ-III ACH were administered after the intervention ended. Additionally, the experimenter gathered the year end DIBELS ORF data from the school to determine each participant’s risk category. The results were used as secondary dependent variables to determine any gains in reading ability/level.

Materials

To implement this study, the experimenter used assessment materials (i.e., WJ-III), the Read Naturally software program, two laptop computers, a digital timer, video-recorder, preferred tangible items as reinforcers, and intervention and data collection materials (i.e., headphones, stories from AIMSweb, data collection sheets, etc.).

Standardized testing materials. The Woodcock-Johnson Test of Achievement – Third Edition (WJ-III ACH; Woodcock et al., 2001) is a highly accurate and valid diagnostic system that has uses for educational, clinical, and research purposes, as well as for diagnosis and guiding educational programming and evaluation. It is a norm-based test that evaluates reading, writing, math, and language. The subtests of the standard battery of the WJ-III ACH that were administered include letter-word identification (LWID), word attack (WA), passage comprehension (PC), and reading fluency (ORF). The median reliability by subtest is 0.94, 0.87, 0.88, and 0.90, respectively. These reliability scores all meet or exceed desired reliability standards. The raw scores and age-
and grade-equivalent scores were used in order to determine the most accurate
information about the reading level of the participants.

*Computers.* Each participant used the RN program that the experimenter installed
on two laptop computers. Each laptop had a pair of headphones that the participants used
to listen to the instructions on the RN program. A Sony Vaio and HP laptop computer
were used for the participants. Both computers ran on the Microsoft XP operating system.

*Read Naturally Software Edition*

The Read Naturally (RN) software program is a computer-assisted reading
program that guides students through an instructional sequence to increase ORF skills.
The experimenter installed the RN program on two computers. Levels 2.0 and 2.5 were
used for this study, as all six students were in second grade. However, none of the
students were reading on grade level at the start of the study. Both levels contained 24
stories each. Students had to progress through the first set of 12 stories in order to
continue to the second set of 12 stories in a level. Tier 1 students progressed through all
of the level 2.0 stories, and read four stories in level 2.5. Time did not permit Tier 2 and 3
students to progress to level 2.5 as the intervention ended at the end of the school year.

The RN stories in levels 2.0 and 2.5 were approximately 95-140 words and
presented five or more key vocabulary words per story. The stand-alone edition of the
RN program was used, which means that the students had to have the disc inserted in the
computer in order to engage with the program. The Flesch-Kincaid Grade Level score in
Word was used to evaluate the readability of eight RN stories. This scale rates text on a
U.S. school grade level. For example, a score of 3.0 means that a third grader can
understand the text. The formula for the Flesch-Kincaid Grade Level score is: 
\[ \text{Flesch-Kincaid} = (.39 \times \text{ASL}) + (11.8 \times \text{ASW}) - 15.59 \]
where:
- ASL = average sentence length (the number of words divided by the number of sentences)
- ASW = average number of syllables per word (the number of syllables divided by the number of words)

The eight RN baseline probe grade level scores ranged from 1.7 to 4.0, according to the above formula. An example of a RN story and comprehension questions are given in Appendix A.

Generalization passages. There were two generalization assessments. The first consisted of stories with traditional narrative texts taken from AIMSweb (Pearson, 2008) progress monitoring probes. The passages were selected from the second grade level material. The participants received one of these passages for every two treatment probes during baseline and throughout the study. Flesch-Kincaid Grade Level scores on some selected AIMSweb passages ranged from 1.0 to 2.2. The second generalization assessment was taken from the participant’s classroom reading material (Direct Instruction - Reading Mastery; Engelmann & Bruner, 1997). The participants read a total of five of these passages over the course of the study. Flesch-Kincaid Grade Level scores on the DI passages ranged from 0.0 to 1.9. See Appendices B and C for examples of both generalization assessments.
**Timer.** A digital kitchen timer was used to time the one-minute timings and treatment probes. The countdown timer was set for one minute and the experimenter started the timer upon the participant beginning to read or retell the passage.

**Video recorder.** A Sony digital video-recorder was used to record sessions in each phase for the purpose of procedural integrity and interobserver agreement (IOA).

**Counter.** In order to determine accurately the number of words retold by each participant, the experimenter used a golf stroke counter. Each time the participant retold a word, the experimenter pressed down on the counter. After the participant finished retelling the story, the number on the counter was recorded on the data sheet.

**Data collection sheets.** The experimenter used a data collection sheet to record the number of correct and incorrect words read for each participant in every session, as well as comprehension information. The data collection sheet consisted of columns and rows for recording such information for each student. See Appendix D for an example.

**Reinforcers.** Each student was given a chart that contained a space for every day of the week. The students were given stickers on the chart for exhibiting appropriate behavior and academic performance, which served as visual reinforcement. Preferred tangible reinforcers were given to the participants upon reaching a predetermined level of stickers on their chart.

**Interobserver Agreement**

A second observer independently recorded data for 30% of pre- and post-tests, baseline, training, intervention, and generalization probes. The second observer listened to the recordings and scored independently the CWPM, IWPM, and oral retell for the
treatment probes. The total number of CWPM, IWPM, and oral retell was calculated by the primary investigator and the second observer, independently, on their own copies of the stories. When there were disagreements, they listened to the recording and determined whether or not the word in question was read correctly.

Interobserver agreement was assessed using the exact agreement method. This is the most rigorous method of calculating IOA for event data (Cooper et al., 2007). With this method, both observers must record the exact number of behaviors per session. For the purposes of this study, the behaviors were the number of words read correctly, incorrectly, and retold in one-minute. Agreement was calculated using the following formula: Number of Agreements/ (Number of Agreements + Number of Disagreements) x 100 = ____%. Interobserver agreement was calculated separately for each story passage and then averaged for all of the passages across all of the participants. Original IOA ranged from 75-100%, with an average of 86%. Because there was a permanent product (recordings) of the probe sessions, there was 100% agreement between both observers. Instances when the observers disagreed on the number of CWPM, IWPM, or number of words retold, the observers listened to the recording to determine whether or not to count the word(s) in question.

Procedural Integrity

Procedural integrity data were collected on 33% of the sessions during training and intervention. The second observer listened to the recordings and completed a checklist to ensure that the each component of the intervention was being implemented completely and accurately. Procedural integrity was established by calculating the
percentage of treatment steps implemented correctly during the session, and averaged across the total number of sessions. The steps were implemented correctly 100% of the time on training and data collection sessions. See Appendices E and G for an example of each checklist. Procedural integrity was also assessed on the correct use of the program by the participants, and on the actions of the RN program. Correct use of the program by the participants ranged from 75% to 100%, with an average of 93%. The RN program functioned properly from 91% to 100% of the time, with an average of 99%. See Appendices F and H for examples of the checklists.

Social Validity

The satisfaction of the participants and teachers was assessed by obtaining measures of social validity. There was a separate measure for the participants and the teachers, and the questions were generated by the experimenter. The measures were administered at the conclusion of the study. The participants dictated his/her answers to the experimenter in the clinical setting; the teachers filled out the questionnaires in private and returned them to the experimenter within three days.

The student questionnaire asked five questions regarding the RN intervention. All six participants answered “yes” or “no” to each question and then told why they responded as such. The teacher questionnaire also asked five questions regarding the RN intervention. Four of the five questions asked teachers to respond to multiple-choice questions on a Likert-type scale; one question asked teachers to respond “yes” or “no.” There was also ample space on the bottom of the questionnaire for additional comments.
Both teachers returned the social validity questionnaire with comments. See Appendices I and J for examples of both questionnaires.
Chapter 4: Results

This chapter presents the results of this study for all of the dependent variables. The main dependent variables were ORF, as measured in CWPM; and comprehension, as measured by the percentage of comprehension questions answered correctly and the number of words re-told per reading passage on RN, AIMSweb, and DI stories. The second set of dependent variables was the pre- and post-test scores on the WJ-III and DIBELS. Social validity data for the teachers and participants are also reported.

Primary Dependent Variables

Following are the results obtained across all participants over the course of the Read Naturally intervention. Specifically, results were analyzed by the participants’ ORF and comprehension on treatment stories (RN), and two types of generalization passages (AIMSweb and DI). Using the multiple-baseline across participants design allowed for a comparison of baseline versus treatment. Embedding changing criterion tactics allowed for analyzing the impact of increasing goals on the participants’ ORF.
Effect Sizes

Cohen’s $d$ is one method that can be used to measure effect sizes. For the purpose of this study, effect sizes were calculated for CWPM on RN treatment passages, and both measures of generalization (AIMSweb & DI). The following formula was used to calculate effect sizes:

$$d = \frac{M_1 - M_2}{\sqrt{S_p^2}}$$

Baseline and treatment were considered two groups in this method of calculating Cohen’s $d$. Thus, Cohen’s $d$ was calculated by subtracting the baseline average from the treatment average and dividing that number by the pooled standard deviation. The standard deviations were pooled due to the different number of probes in baseline versus treatment. Gravetter and Wallnau (2007) note that according to Cohen (1988), if the magnitude of $d$ falls between 0-0.2, it is a small effect. A medium effect is when $d$ is between 0.2-0.8, and a large effect is 0.8 or higher.

North River Academy

Figure 4.1 presents the multiple baseline ORF data for all six students. This figure shows the number of CWPM and IWPM on Read Naturally (RN) treatment probes, and the number of CWPM on AIMSweb generalization passages in the treatment setting, and Direct Instruction (DI) generalization passages in the classroom setting. Table 4.1
presents these average and corresponding effect sizes and standard deviations. Table 4.2 presents the criterion changes in CWPM for each participant, and the number of stories that he/she completed in each criterion change. Figure 4.2 presents the retell data for all six participants, and Table 4.3 presents the average words re-told for each participant in baseline and treatment across RN, AIMS, and DI passages. Figure 4.3 presents the comprehension data for all six participants, and Table 4.4 presents the average percentage of comprehension questions answered correctly for each participant in baseline and treatment across the RN, AIMS, and DI probes. Following are the results by participant.

Tabitha

Tabitha, the first participant, averaged 20 CWPM on RN stories in baseline and 65 CWPM during the intervention, which is a very large effect size of 3.02. She had an initial goal of 24 CWPM in her first criterion, and a final goal of 84 CWPM in her sixth criterion change. This is an increase of 60 CWPM over the course of the intervention. Tabitha’s data path for RN stories shows a variable trend, but it began to stabilize toward the end of the intervention. She also showed a generalization increase in the average CWPM on AIMSweb passages from baseline to intervention, 19 to 30 CWPM, respectively. That was a large effect size of 1.44. She had an initial upward trend on these AIMSweb passages, followed by a decrease, then another increase toward the end of the intervention. On the second measure of generalization, DI passages, Tabitha increased from 48 to 59 CWPM, which also reveals a quite large effect size of 2.81.

Tabitha also showed increases in the average percentage of comprehension questions answered correctly; however, these numbers were not as robust. On the RN
Even though Tabitha’s retell data on RN probes were variable, there was an increase in average number of words re-told during baseline, 14; to 39 during the intervention. This was over double the number of average words re-told during baseline, and there were no overlapping data points between baseline and treatment. Her data path for the first measure of generalization, AIMSweb passages, revealed a similar trend. There was one outlier in baseline, but her average number of words re-told from baseline to intervention increased from 16 to 24. Tabitha made greater gains in words re-told per minute on the second measure of generalization, DI. Initially, she re-told an average of 19 words per minute during baseline, in comparison to 39 words per minute during intervention. Again, she doubled the average number of words re-told per minute. Her data path for DI passages showed a generally increasing trend.

**Diana**

Diana averaged 19 CWPM on RN stories in baseline and 53 CWPM during the intervention, which is a very large effect size of 2.48. She had an initial goal of 23 CWPM in her first criterion, and a final goal of 72 CWPM in her sixth criterion change. This is an increase of 49 CWPM over the course of the intervention. Diana’s data path for RN stories shows a slow, upward trend. She also showed a generalization increase in the average CWPM on AIMSweb passages from baseline to intervention, 11 to 21 CWPM,
respectively. That was a large effect size of 2.81. Her data path on these passages remained quite low and stable throughout the intervention. On the second generalization measure, DI passages, Diana increased from 32 to 39 CWPM, which also reveals a slightly large effect size of 0.98.

Diana did not fare as well on the comprehension probes. On the RN stories, Diana answered 30% correct during baseline versus 26% in treatment. Her comprehension scores on the AIMSweb questions also decreased from an average of 47% correct in baseline to 34% correct during treatment. She showed similar decreases on the DI stories, from 60% correct in baseline to only 45% correct in treatment.

However, Diana did show improvement on another measure of comprehension; the number of words re-told per story. Her retell data on RN probes were variable; but there was an increase in average number of words re-told during baseline, 9; to 32 during the intervention. This was over three and a half times the number of average words re-told during baseline, and there were no overlapping data points between baseline and treatment. Her data path for the first measure of generalization, AIMSweb passages, revealed a slightly increasing trend with an average number of words re-told from baseline to intervention increasing from 7 to 22. Diana scored similarly on the second measure of generalization, DI, with a re-tell score of 7 words per minute during baseline and 29 words per minute during intervention. Again, she more than tripled the average number of words re-told per minute.
Javon

Javon, the third participant, averaged 23 CWPM on RN stories in baseline and 68 CWPM during the intervention, which is a very large effect size of 3.61. He had an initial goal of 33 CWPM in his first criterion, and a final goal of 84 CWPM in his sixth criterion change. This was an increase of 51 CWPM over the course of the intervention. Javon’s data path for RN stories shows a slowly increasing trend with only one overlapping data point. He also showed a generalization increase in the average CWPM on AIMSweb passages from baseline to intervention, 19 to 29 CWPM, respectively. That was a large effect size of 2.11. He had a stable data path on the AIMSweb passages. On the DI generalization passages, Javon increased from 29 to 52 CWPM, which is also a large effect size of 1.09.

Javon’s scores on comprehension questions reveals mixed results. Javon only made increases on the RN stories, answering 26% correct during baseline and 40% in treatment. His comprehension scores on the AIMSweb questions remained the same at 53% during baseline and treatment. However, his scores on the DI stories decreased from 80% during baseline to 35% during intervention. Possibilities for this decrease will be discussed in Chapter five.

Javon’s retell scores improved across all measures. The data path for the RN probes began with a decreasing trend, however, these data stabilized with an upward trend by the end of the intervention. The average number of words re-told on RN stories increased from 14 during baseline to 38 during the intervention. This was over two and a half times the number of average words re-told during baseline with no overlapping data
points. His data path for AIMSweb passages revealed a slowly increasing trend as the average number of words re-told from baseline to intervention increased from 18 to 26. On the DI passages, Javon re-told 24 words per minute during baseline and an average of 33 words per minute during intervention. His data path for DI passages showed an increasing trend throughout.

**Angela**

The fourth participant, Angela, averaged 38 CWPM on RN stories in baseline and 95 CWPM during the intervention, which is a very large effect size of 3.15. She had an initial goal of 50 CWPM in her first criterion, and a final goal of 113 CWPM in her sixth criterion change. This was an increase of 63 CWPM over the course of the intervention. Angela’s data path for RN stories shows a somewhat variable yet upward trend. She also showed a generalization increase in the average CWPM on AIMSweb passages from baseline to intervention, 42 to 65 CWPM, respectively. That was a large effect size of 1.56. She had an initial upward trend on these AIMSweb passages, but it began to decline over the course of the intervention. On the DI generalization passages, Angela increased from 60 to 81 CWPM, which also reveals very robust effect size of 6.15.

The average percentage of comprehension questions answered correctly by Angela also increased over the course of the study. On the RN stories, Angela answered 31% correct during baseline and 51% correct in treatment. Her best gains in comprehension scores were on the AIMSweb passages, with an average of 47% correct in baseline to 75% correct during treatment. Her scores on the DI stories went from 60% correct in baseline to 70% correct in treatment.
Angela’s retell data for the RN probes were quite variable during the intervention. She did have an increase in average number of words re-told from baseline to intervention, 10 to 20, respectively. However, her data path on AIMSweb passages revealed a constant upward trend, averaging only six words re-told per minute during baseline to 20 words during intervention. She also increased the average number of words re-told on DI passages, from 16 words per minute during baseline to 27 words per minute during intervention. Her data path for DI passages showed a generally increasing trend.

Quantasia

Quantasia, the fifth participant, averaged 32 CWPM on RN stories in baseline and 71 CWPM during the intervention, which is a very large effect size of 2.67. She had an initial goal of 41 CWPM in her first criterion, and a final goal of 92 CWPM in her fifth criterion change. This was an increase of 51 CWPM over the course of the intervention. Quantasia’s data path for RN stories shows a generally increasing trend over the course of the intervention. She also showed a generalization increase in the average CWPM on AIMSweb passages from baseline to intervention, 36 to 46 CWPM, respectively. That was a slightly large effect size of 0.99. Her data path on these AIMSweb passages revealed a variable trend, with three overlapping data points. On the DI generalization passages, Quantasia increased from an average of 56 to 68 CWPM, with a medium effect size of 0.79.

Quantasia’s scores for comprehension questions answered correctly revealed mixed results. She only increased the number of comprehension questions answered correctly on RN probes, from 28% correct during baseline to 52% correct in treatment.
Even though Quantasia’s retell data on RN probes were variable, there was an increase in average number of words re-told from 18 in baseline to 26 during intervention. Her data path for AIMSweb passages revealed a similar variable trend with the average number of words re-told from baseline to intervention increasing from 15 to 28. Quantasia also had an increase on DI passages, from an average of 20 words per minute during baseline to 34 words per minute during intervention. This was her strongest data path, which showed a continuous upward trend.

*Tamika*

The last participant, Tamika, averaged 45 CWPM on RN stories in baseline and 89 CWPM during the intervention, which is a very large effect size of 3.86. She had an initial goal of 48 CWPM in her first criterion, and a final goal of 109 CWPM in her fifth criterion change. This was an increase of 61 CWPM over the course of the intervention. Tamika’s data path for RN stories shows an increasing variable trend. She also showed a generalization increase in the average CWPM on AIMSweb passages from baseline to intervention, 50 to 67 CWPM, respectively. That was a very large effect size of 2.35. Her AIMSweb data path showed a similar increasing variable trend with no overlapping data points. On the DI generalization passages, Tamika actually decreased from an average of 82 to 79 CWPM, which reveals a negative effect size of -0.29.
On the RN stories, Tamika answered 48% of the comprehension questions correctly during baseline versus 47% in treatment. She did, however, show increases in the percentage of comprehension questions answered correctly on both measures of generalization. Her comprehension scores on the AIMSweb questions increased from an average of 70% correct in baseline to 83% correct during treatment. Her best gains in comprehension were on the DI stories, from 30% in baseline to 73% in treatment.

Tamika’s retell data were variable for RN and AIMSweb passages. On the RN probes, Tamika re-told an average of 17 words per minute in baseline and 36 words during intervention. On the AIMSweb passages, she retold an average of 24 and 32 words per minute, respectively. Her data path for DI passages showed an increasing trend during the intervention. Tamika re-told an average of 26 words per minute during baseline and an average of 35 words during intervention.
Figure 4.1 Treatment and generalization ORF data
Figure 4.1 continued

![Graphs showing baseline and treatment performance for Angela, Quantasia, Probes, and Tamika. The graphs depict Words per Minute over time.](image-url)
<table>
<thead>
<tr>
<th>Student</th>
<th>RN Probes</th>
<th>AIMS – Gen. 1 Probes</th>
<th>DI – Gen. 2 Probes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BL</td>
<td>TX</td>
<td>ES (SD)</td>
</tr>
<tr>
<td>Tabitha</td>
<td>20</td>
<td>65</td>
<td>3.02 (14.80)</td>
</tr>
<tr>
<td>Diana</td>
<td>19</td>
<td>53</td>
<td>2.48 (13.49)</td>
</tr>
<tr>
<td>Javon</td>
<td>23</td>
<td>68</td>
<td>3.61 (12.37)</td>
</tr>
<tr>
<td>Angela</td>
<td>38</td>
<td>95</td>
<td>3.15 (18.08)</td>
</tr>
<tr>
<td>Quantasia</td>
<td>32</td>
<td>71</td>
<td>2.67 (14.68)</td>
</tr>
<tr>
<td>Tamika</td>
<td>45</td>
<td>89</td>
<td>3.86 (11.41)</td>
</tr>
</tbody>
</table>

Table 4.1 Baseline and treatment CWPM averages; effect sizes and pooled standard deviations included
<table>
<thead>
<tr>
<th>Student</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabitha</td>
<td>24</td>
<td>43</td>
<td>58</td>
<td>67</td>
<td>71</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(3)</td>
<td>(3)</td>
<td>(3)</td>
<td>(6)</td>
<td>(2)</td>
</tr>
<tr>
<td>Diana</td>
<td>23</td>
<td>36</td>
<td>44</td>
<td>55</td>
<td>61</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(3)</td>
<td>(3)</td>
<td>(3)</td>
<td>(6)</td>
<td>(2)</td>
</tr>
<tr>
<td>Javon</td>
<td>33</td>
<td>54</td>
<td>60</td>
<td>71</td>
<td>82</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(2)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(2)</td>
</tr>
<tr>
<td>Angela</td>
<td>50</td>
<td>72</td>
<td>85</td>
<td>100</td>
<td>107</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(2)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(2)</td>
</tr>
<tr>
<td>Quantasia</td>
<td>41</td>
<td>45</td>
<td>68</td>
<td>82</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(1)</td>
<td></td>
</tr>
<tr>
<td>Tamika</td>
<td>48</td>
<td>80</td>
<td>89</td>
<td>102</td>
<td>109</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(1)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2 Criterion changes in CWPM; (number of stories per criterion) included
Figure 4.2 Treatment and generalization retell data
Figure 4.2 continued

![Graph showing the words re-told per minute for Baseline and Treatment phases for Angela, Quantasia, and Tamika across different probes.](image)

Words Re-Told per Minute

Probes

Angela

Quantasia

Tamika
<table>
<thead>
<tr>
<th>Student</th>
<th>BL</th>
<th>TX</th>
<th>BL</th>
<th>TX</th>
<th>BL</th>
<th>TX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabitha</td>
<td>14</td>
<td>39</td>
<td>16</td>
<td>24</td>
<td>19</td>
<td>39</td>
</tr>
<tr>
<td>Diana</td>
<td>9</td>
<td>32</td>
<td>7</td>
<td>22</td>
<td>7</td>
<td>29</td>
</tr>
<tr>
<td>Javon</td>
<td>14</td>
<td>38</td>
<td>18</td>
<td>26</td>
<td>24</td>
<td>33</td>
</tr>
<tr>
<td>Angela</td>
<td>10</td>
<td>20</td>
<td>6</td>
<td>20</td>
<td>16</td>
<td>27</td>
</tr>
<tr>
<td>Quantasia</td>
<td>18</td>
<td>26</td>
<td>15</td>
<td>28</td>
<td>20</td>
<td>34</td>
</tr>
<tr>
<td>Tamika</td>
<td>17</td>
<td>36</td>
<td>24</td>
<td>32</td>
<td>26</td>
<td>35</td>
</tr>
</tbody>
</table>

Table 4.3 Baseline and treatment average words re-told per minute
Figure 4.3 Treatment and generalization comprehension data
Figure 4.3 continued

![Bar chart showing the percentage of correct comprehension questions for Angela, Quantasia, and Tamika during baseline and treatment phases across different probes.](image)

- **Baseline** vs. **Treatment**
- **Probes** range from 1 to 27
- **Percentage of Correct Comprehension Questions** range from 0% to 100%

- **Angela**
- **Quantasia**
- **Tamika**

120
<table>
<thead>
<tr>
<th>Student</th>
<th>RN Probes</th>
<th>AIMS – Gen. 1 Probes</th>
<th>DI – Gen. 2 Probes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BL</td>
<td>TX</td>
<td>BL</td>
</tr>
<tr>
<td>Tabitha</td>
<td>40</td>
<td>44</td>
<td>20</td>
</tr>
<tr>
<td>Diana</td>
<td>30</td>
<td>26</td>
<td>47</td>
</tr>
<tr>
<td>Javon</td>
<td>26</td>
<td>40</td>
<td>53</td>
</tr>
<tr>
<td>Angela</td>
<td>31</td>
<td>51</td>
<td>47</td>
</tr>
<tr>
<td>Quantasia</td>
<td>28</td>
<td>52</td>
<td>60</td>
</tr>
<tr>
<td>Tamika</td>
<td>48</td>
<td>47</td>
<td>70</td>
</tr>
</tbody>
</table>

Table 4.4 Baseline and treatment average percentage of comprehension questions answered correctly
Secondary Dependent Variables

The results of the pre- and post-tests on DIBELS and WJ-III represent the secondary dependent variables in this study. The experimenter assessed each participant on the letter-word identification (LWID), word attack (WA), reading fluency (RF), and passage comprehension (PC) sub-tests of the WJ-III prior to commencing and at the conclusion of the study. Each participant’s raw scores were used to compute grade equivalence, standard scores, and classifications according to the WJ-III. The experimenter gathered the winter and spring DIBELS benchmark data from the participants’ teacher. The DIBELS benchmark assessment for ORF reports scores using CWPM and risk categories. These categories are: at-risk, some risk, and low risk.

Table 4.5 presents the scores for the WJ-III pre- and post-test scores for LWID and WA for all six participants. Table 4.6 presents the scores for the WJ-III pre- and post-test scores for RF and PC for all six participants. Table 4.7 presents the scores and risk category for each participant on the winter and spring DIBELS benchmark assessments.

Tabitha

Tabitha’s grade equivalence scores increased from pre- to post-test for all WJ-III subtests. On the LWID subtest, her scores increased from 1.5 to 1.9, resulting in a low average classification. Her WA scores increased from 1.5 to 2.0, with a low average to average classification. Reading fluency scores increased slightly from pre- to post-test, 1.5 and 1.7, respectively. This classified her in the low to low average range. The most
gains were made in PC, from 1.2 to 1.9, which changed her classification from very low to low average. Tabitha also gained 23 CWPM on DIBELS, scoring 17 CWPM in the winter and 40 CWPM in spring. These scores still qualify as at-risk.

_Diana_

Diana’s grade equivalence scores are commensurate with her scores on the primary dependent variables. She actually regressed on the LWID and RF subtests. Her grade equivalence of 1.7 remained the same from pre- to post-test in LWID, however, due to her standard score, her classification decreased from low average on the pre-test to low to low average. Likewise, a 1.8 (pre) and 1.4 (post) grade equivalence score on the RF subtest decreased her classification from low average to average on the pre-test to low to low average at post-test. She did, however, make gains from pre- to post-test on WA, from 1.8 to 2.3, respectively. This classified her in the average range. Even though her classification remained the same on PC, her grade equivalence score increased from 1.3 to 1.5. Diana remained very low on DIBELS, scoring just 18 CWPM in winter and 22 on the spring assessment. She remained at-risk for reading failure according to DIBELS.

_Javon_

Javon’s pre- and post-test scores for LWID increased from 1.6 to 2.1, respectively. This resulted in a low average classification. An increase from 1.5 to 1.7 on the WA subtest resulted in a low average to average classification. His grade equivalence scores on RF resulted in a low to low average classification (1.9), making gains from very low to low average on the pre-test (<K.0). Javon also made gains in PC, from 1.2 (pre) to 1.8 (post), resulting in low to low average classification. According to DIBELS, he remained
at-risk, scoring 18 CWPM in the winter and 32 CWPM in the spring. This was an increase of 14 CWPM over the course of the study.

_Angela_

Angela’s increases in grade equivalence scores resulted in an increase in classification on every WJ-III subtest. On LWID, she scored 1.8 (pre) and 2.7 (post), resulting in an average classification. Her post-test score of 1.7 on the WA subtest resulted in a low average to average classification, which was higher than her 1.3 score on the pre-test. She scored in the average range in reading fluency, with a pre- and post-test score of 1.5 and 2.5, respectively. She also made gains in PC, from 1.7 on the pre-test to 2.3 at post-test, classifying her in the low average to average range. Angela made the most progress of any participant on DIBELS. She read 30 CWPM on the winter assessment, placing her in the at-risk category. On the spring assessment, Angela read 81 CWPM, placing her in the some-risk category. This was an increase of 51 CWPM over the course of the study. Furthermore, she was only nine words from being in the low-risk category according to DIBELS.

_Quantasia_

Pre- (2.1) and post-test (2.4) grade equivalence scores on the LWID subtest for Quantasia resulted in a classification of low average to average. She had this same classification on the WA and RF subtests. She scored 1.8 (pre) and 1.9 (post) on WA and 1.6 (pre) and 2.2 (post) on RF. Her scores on PC increased from 1.6 to 1.7, resulting in a low to low average classification. Quantasia gained 18 CWPM from winter to spring on
DIBELS, reading 25 and 43 words, respectively. This still classifies her in the at-risk category according to DIBELS.

Tamika

Tamika made gains in grade equivalence scores in three of the four subtests on the WJ-III. Her scores on the LWID subtest increased from 2.9 to 3.0, resulting in an average classification. Even though her grade equivalence scores on the WA subtest decreased from 2.9 to 2.7, these scores still classified her as average. She also had an average classification in reading fluency, with an increase from 2.4 (pre) to 2.6 (post). On the PC subtest, she scored 1.8 and 2.3 on the pre- and post-test, respectively. This resulted in a low average to average classification. Tamika read 23 CWPM on the DIBELS benchmark assessment in the winter and 66 CWPM in the spring. Even though this resulted in a 23 CWPM increase over the course of the study, these scores still placed her in the at-risk category.
<table>
<thead>
<tr>
<th>Students</th>
<th>1GE</th>
<th>2SS</th>
<th>3Class</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GE</td>
<td>SS</td>
<td>Class</td>
<td>GE</td>
<td>SS</td>
<td>Class</td>
<td>GE</td>
</tr>
<tr>
<td>Tabitha</td>
<td>1.5</td>
<td>77-83</td>
<td>Low-</td>
<td>1.9</td>
<td>80-86</td>
<td>L.Ave</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>L.Ave</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diana</td>
<td>1.7</td>
<td>83-88</td>
<td>L.Ave</td>
<td>1.7</td>
<td>77-82</td>
<td>Low-</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L.Ave</td>
<td></td>
</tr>
<tr>
<td>Javon</td>
<td>1.6</td>
<td>81-86</td>
<td>L.Ave</td>
<td>2.1</td>
<td>84-89</td>
<td>L.Ave</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angela</td>
<td>1.8</td>
<td>84-90</td>
<td>L.Ave-Ave</td>
<td>2.7</td>
<td>94-99</td>
<td>Ave</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantasia</td>
<td>2.1</td>
<td>90-95</td>
<td>Ave</td>
<td>2.4</td>
<td>89-94</td>
<td>L.Ave-Ave</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamika</td>
<td>2.9</td>
<td>103-108</td>
<td>Ave</td>
<td>3.0</td>
<td>99-104</td>
<td>Ave</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 GE = Grade Equivalent Scores  
2 SS = Standard Score 68% Band  
3 Class = Classification based on SS  
4 L.Ave = Low Average  

Table: 4.5 Woodcock Johnson III pre-test and post-test scores for LWID and WA sub-tests
<table>
<thead>
<tr>
<th>Students</th>
<th>GE</th>
<th>SS</th>
<th>Class</th>
<th>GE</th>
<th>SS</th>
<th>Class</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabitha</td>
<td>1.5</td>
<td>75-89</td>
<td>Low-4L.Ave</td>
<td>1.7</td>
<td>75-86</td>
<td>Low-L.Ave</td>
<td>1.2</td>
<td>65-75</td>
<td>5V.Low</td>
<td>1.9</td>
</tr>
<tr>
<td>Diana</td>
<td>1.8</td>
<td>82-94</td>
<td>L.Ave-Ave</td>
<td>1.4</td>
<td>70-83</td>
<td>Low-L.Ave</td>
<td>1.3</td>
<td>68-78</td>
<td>V.Low-Low</td>
<td>1.5</td>
</tr>
<tr>
<td>Javon</td>
<td>&lt;K.0</td>
<td>59-86</td>
<td>V.Low-L.Ave</td>
<td>1.9</td>
<td>79-89</td>
<td>Low-L.Ave</td>
<td>1.2</td>
<td>65-75</td>
<td>V.Low-Low</td>
<td>1.8</td>
</tr>
<tr>
<td>Angela</td>
<td>1.5</td>
<td>75-89</td>
<td>Low-L.Ave</td>
<td>2.5</td>
<td>90-100</td>
<td>Ave</td>
<td>1.7</td>
<td>80-89</td>
<td>L.Ave</td>
<td>2.3</td>
</tr>
<tr>
<td>Quantasia</td>
<td>1.6</td>
<td>78-91</td>
<td>Low-Ave</td>
<td>2.2</td>
<td>85-95</td>
<td>L.Ave-Ave</td>
<td>1.6</td>
<td>78-86</td>
<td>Low-L.Ave</td>
<td>1.7</td>
</tr>
<tr>
<td>Tamika</td>
<td>2.4</td>
<td>93-103</td>
<td>Ave</td>
<td>2.6</td>
<td>91-101</td>
<td>Ave</td>
<td>1.8</td>
<td>83-92</td>
<td>L.Ave-Ave</td>
<td>2.3</td>
</tr>
</tbody>
</table>

1 GE = Grade Equivalent Scores  
2 SS = Standard Score 68% Band  
3 Class = Classification based on SS  
4 L.Ave = Low Average  
5 V.Low = Very Low

Table: 4.6 Woodcock Johnson III pre-test and post-test scores for RF and PC sub-tests
<table>
<thead>
<tr>
<th>Student</th>
<th>2nd Grade Benchmark Scores</th>
<th>2nd Grade Benchmark Scores</th>
<th>2nd Grade Benchmark Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DIBELS Winter</td>
<td>DIBELS Spring</td>
<td>DIBELS Winter</td>
</tr>
<tr>
<td>Student</td>
<td>ORF</td>
<td>Risk Status</td>
<td>ORF</td>
</tr>
<tr>
<td>Tabitha</td>
<td>17</td>
<td>At-Risk</td>
<td>40</td>
</tr>
<tr>
<td>Diana</td>
<td>18</td>
<td>At-Risk</td>
<td>22</td>
</tr>
<tr>
<td>Javon</td>
<td>18</td>
<td>At-Risk</td>
<td>32</td>
</tr>
<tr>
<td>Angela</td>
<td>30</td>
<td>At-Risk</td>
<td>81</td>
</tr>
<tr>
<td>Quantasia</td>
<td>25</td>
<td>At-Risk</td>
<td>43</td>
</tr>
<tr>
<td>Tamika</td>
<td>43</td>
<td>At-Risk</td>
<td>66</td>
</tr>
</tbody>
</table>

Table 4.7 Winter and spring DIBELS ORF benchmark scores and risk status
Results: Social Validity

All of the participants and teachers completed social validity questionnaires. Results are presented separately for the teachers and the participants. The teacher and student questionnaires can be found in Appendices I and J, respectively.

Teachers

Both teachers felt that all of the students’ overall reading had improved over the course of the intervention, and would say that the students became more fluent readers. However, Teacher 2 noted that the students who put forth more effort became better readers. The teachers agreed that all of the students enjoyed the intervention; they would allow their students to participate in similar interventions in the future, and thought that reading interventions are very important for struggling readers. Furthermore, Teacher 2 commented that her students’ confidence had grown as a result of this intervention, the students tried harder in class, and would benefit even more if the intervention could have lasted longer. She noted that the program was very helpful and that her students would become even better readers if they continued using this program.

Participants

The students answered questions about their motivation, enjoyment of the computer program, and their reading abilities. There were five questions, all of which were “yes” or “no” answers and included an open-ended “Why?” response. All six of the participants responded “yes” to each question. This means that every student liked reading on the computer and would like to continue reading on the computer in the future and thought that he/she became a better reader. Every student said that they liked
working for prizes, but somewhat surprisingly, every student also responded that they
would have worked on the computer even if they had not received prizes.

The responses to the open-ended question “Why?” were quite telling. These
struggling readers were willing to work on this computer program even if no prizes were
given because, “I want to know how to read” (Tabitha) and “I want to beat my goal”
(Angela). Two more students shared the same sentiment. Tamika would have worked on
the computer without receiving prizes simply because she liked using the computer
technology. Javon also agreed that reading on the computer was “…fun and very
interesting.” Four students mentioned beating their goal as part of what they enjoyed
about the program, and five students liked how the computer program helped them learn
new words and practice reading a lot. Anecdotally, five of the participants were sad that
the intervention was over and that they could not work on the computer anymore.
Chapter 5: Discussion

This study examined the effectiveness of a computer-assisted reading program, Read Naturally (RN), on the oral reading fluency (ORF), comprehension, and generalization of second graders who were at risk for reading failure. Six students participated in this study over the course of four months. The students received the intervention (RN) three to four times a week for approximately 7 to 12 weeks. All six students were trained to use the RN program, and were able to advance through the sequence of instructional activities with little to no assistance.

A multiple-baseline across participants design with embedded changing criterion tactics was used in this study. This design allowed the experimenter to note the effects of increasing the goal criterion (in CWPM) on the participants’ ORF. Two students were in each tier. Tier one began intervention after a stable baseline of six data points. After a treatment effect was shown for tier one in their first criterion change, tier two received one more baseline probe before beginning the intervention. Tier three students also received one more baseline probe at this same time. After a treatment effect was shown for tier two in their first criterion change, tier three was administered one more baseline probe before entering into intervention. Once all three tiers had entered the intervention, the experimenter increased each participant’s goal (CWPM) based on individual performance. Tiers one and two had six criterion changes over the course of the study, and tier three had five criterion changes. These data can be seen in Table 4.2.
Throughout the study, data were taken on various dependent variables. The experimenter collected data on the correct words per minute (CWPM) and incorrect words per minute (IWPM) read by each participant on the RN passages as well as on two measures of generalization (AIMSweb and DI). One AIMSweb passage was given for every two RN stories. Depending on the tier, there were either one or two DI passages given during baseline and three or four DI passages during intervention. Each participant received five DI probes; either 23 or 26 RN probes; and 11 or 13 AIMSweb probes over the course of the study. Additionally, data was collected on the number of comprehension questions answered correctly by each participant on all three types of stories (i.e., RN, AIMSweb, and DI); as well as on the number of words re-told in a minute.

Pre- and post-test measures were given individually to each participant approximately one week before the study began and during the last week of the study. These consisted of the letter-word identification (LWID), word attack (WA), reading fluency (RF), and passage comprehension (PC) subtests of the WJ-III. The experimenter also gathered DIBELS ORF winter and spring benchmark data from the classroom teacher.

This chapter will discuss study results according to each research question. Along with an analysis of each research question, limitations of the study, practical implications, and future research will be discussed.

Research Question One

What effect will the RN computer software program have on the oral reading fluency of urban 2nd-grade students at risk for reading failure?
Correct Words per Minute

The data revealed that every participant’s scores on CWPM on RN probes increased from baseline to intervention, suggesting that this computer program was effective in increasing their ORF. Collectively, the participants gained an average of 44 CWPM from baseline to intervention. Individual gains in CWPM correspond with the collective gains. Angela’s average individual gains at 57 CWPM were actually higher than the group average. This is particularly impressive since Angela was an English language learner, suggesting that this instructional program may be equally effective with ELLs as well as native English speakers. This might be due to the format of RN, as the students interact with the program and are rarely passively receiving information. Gersten and Baker (2000) also point out that ELLs made increases in reading when they were actively responding to their schoolwork.

The increase in CWPM for the participants in the current study are consistent with the literature on the success of repeated reading interventions (Begeny et al., 2006; Chard et al., 2002; Hitchcock et al.; Lane et al., 2007; Martens et al., 2007; Nelson et al., 2004; Vadasy et al., 2005). More specifically, these results actually support the use of the RN program. Hasbrouck et al. (1999) used the RN program over six years with 214 second- and third-grade students. The intervention lasted approximately 32 weeks, and the participants’ average CWPM increased from the fall to spring of each year. The authors reported the average CWPM gain by week instead of the average of the whole intervention. They note that “typical” students should gain 1.5 CWPM per week (see Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Good & Kaminski, 2002), and their second-grade remedial students gained an average of 1.68 CWPM, exceeding the
“typical” average. This is noteworthy because low-performing students are generally thought to progress more slowly than their typical peers, yet they kept up with and even surpassed their typical peers (Hasbrouck et al.).

One difference between the current study and the Hasbrouck et al. (1999) study is that Hasbrouck et al. used the Master Edition of RN and the current investigation used the Software Edition. However, the teaching techniques (i.e., teacher modeling of reading, RR, and progress monitoring) are the same in both programs. Therefore, the at-risk students in the current study performed just as well or better than the remedial students in the Hasbrouck study. For example, all six students in the current investigation scored above the 1.5 CWPM expected weekly growth on RN probes. In fact, five students ranged between an average of 2.1 and 2.9 CWPM/week, and Angela actually averaged 3.6 CWPM/week. Good and Kaminski (2002) actually note that 1.5 is the realistic goal for second grade growth, whereas 2.0 is an ambitious goal for second graders. In the current study, all six students exceeded the ambitious goal for second graders. This is remarkable considering that all of these students were at-risk for reading failure and started the study reading well below grade level.

Incorrect Words per Minute

Results on the number of IWPM were gathered by counting any errors (i.e., substitutions, omissions, additions, etc.) during the one-minute pass timing activity and entering that number into the RN program. All of the participants decreased the number of errors made on RN passages from baseline to intervention. Their collective average of IWPM decreased from eight to three.
Again, this could have been predicted because of the teaching techniques that the RN program employs. The computer provided teacher modeling of fluent reading and error correction (EC). The computer models fluent reading during the read along activity in the RN sequence. This read along activity is essentially the listening passage preview (LPP). The participants listened to the story three times, and then moved on to the practice reading. This practice reading activity is the RR portion, when the student is required to practice reading the story as many times as needed in order to meet his/her predetermined goal. When a student had difficulty with a word, he/she was able to click on the word and have it read to them. This would be considered a form of error correction. These techniques, LPP and error correction, have been demonstrated to be effective when combined with RR (Ardoin et al., 2007; Ardoin et al., 2008; Begeny, 2006; Martens et al., 2007; Nelson et al., 2004; Welsch, 2007).

Nelson and colleagues (2004) did not use LPP, but did employ EC procedures alone and with RR. The four students in this study progressed through a treatment package of EC, EC plus RR, and finally EC plus RR with previously read materials. The EC condition alone did not confer much benefit to the students as far as increasing the CWPM, but IWPM did decline in this condition. The decline in IWPM continued throughout the next two conditions in which RR was implemented and also served to increase the CWPM for all four students. Therefore, providing students with EC during RR activities has an impact not only on decreasing the number of errors, but also increasing the number of CWPM.
**Criterion Changes**

Throughout the study, the experimenter changed each participant’s goal criterion by averaging the two highest probes from the previous criterion. The participants had to meet or exceed their goal for at least two RN probes in order to change their criterion. This was done to determine how increasing the number of CWPM would affect generalization scores. Results for generalization effects on ORF will be discussed in questions four and seven.

The students in tiers one and two had six criterion changes, and the students in tier three had five criterion changes. Over the short time frame of just three months, the students’ goals were increased by 49 to 63 CWPM. This more than doubled the number of CWPM from their first criterion change. By the end of the study, three students’ goals were over the benchmark of 90 plus CWPM. According to DIBELS, second grade students who can read 90 or more words per minute are on track and at low risk for reading failure.

In general, the students met their goals more often than not. The students missed their goal between three and eight times each over the course of the study. Oftentimes, the students would make comments about “how high” they wanted to read. They were referring to the line on the bar graph in the RN program. A defining feature of the program is progress monitoring. At the end of the cold timing, the participants see their score in CWPM, and also see the line for their goal. If the student passes the story during the pass timing, the bar graph shows the scores for cold timing, their goal, and the pass timing. This type of feedback seemed to motivate the participants. In fact, Ardoin et al. (2008) point out that when teachers set goals and give students feedback on their
performance, that can be even more effective than just having the students repeatedly read a passage. Repeated reading without those motivational components might not be as effective. Continued research in this area is warranted.

Research Question Two

*What effect will the RN computer software program have on the number of comprehension questions answered correctly by the participants?*

The instructional sequence in the RN program includes a quiz activity after the participants have completed the practice read activity. The participants had to have met or exceeded their goal (CWPM) in order to progress to the quiz. The quiz contained four multiple-choice questions and one open-ended question about the story. Responses to these five questions on the quiz were used for the comprehension question data. It is important to note, however, that because there were only five questions, the participants could only score between 0-100%, in 20% increments.

The results for comprehension as measured by the percentage of questions answered correctly were not as strong on the RN probes as the scores for ORF. Four of the six participants increased the percentage of correct comprehension questions answered correctly from baseline to intervention. It is important to note, however, that each participant’s performance on this comprehension measure was comparable to their growth in ORF. For example, Diana scored quite low in ORF and actually regressed in comprehension; whereas Angela scored quite high in ORF and was one of the highest performers on RN comprehension.

Some researchers (Hasbrouck et al., 1999, Stevens, 2006) have stated that passage comprehension might increase as readers become more fluent. Stahl (2004) noted
numerous studies that have demonstrated that students who read connected text repeatedly or with assistance show increases in both reading comprehension and fluency measures. Kuhn & Stahl (2003) agree that fluency related instruction increases reading comprehension.

The findings in this study are somewhat contradictory to previous research on the fluency-comprehension link. There are a few possibilities of why the current participants did not fare as well in answering comprehension questions as opposed to retelling the story. One possibility is because there were a limited number of responses that the participants could make (only five comprehension questions). Therefore, these data could be a function of the assessment modality. Markell and Deno (1997) also stated that growth in reading might not be accurate by reporting the percentage correct on comprehension questions. They drew that conclusion because percentages are less sensitive to comprehension change as there are only a limited number of questions per reading passage that students can answer, which reduces the actual percentage scores possible.

Another possibility for low scores in the percentage of comprehension questions answered correctly could be that the participants might have misinterpreted the question or misread some of the words. The participants were required to read the questions and answers themselves, so if they struggled with certain words during the practice reading, they might not have to know how to say/read those words when they appeared in the comprehension questions. In a study by Markell and Deno (1997), the experimenters actually read the comprehension questions aloud to the participants and the participants orally stated their answers. In general, the participants averaged over 50% correct. Even
though a student answering 50% of comprehension questions correctly might not be considered to understand the text, it is still higher than most of the participants in the current study. For example, two participants’ comprehension scores actually decreased from baseline to intervention; and only two students’ scores (Angela and Quantasia) exceeded 50%, answering 51% and 52% correct, respectively.

Yet another possibility for the participants’ level of responding on comprehension questions is that the experimenter did not directly teach the participants strategies on how to choose the correct answer. For example, the first question always asked, “What is most of this story about?” That is a question about the main idea of the story. If the participants did not understand the concept of the main idea, then he/she might choose the wrong answer, like one of the literal fact distracters. Researchers (Kamil, 2004; Lyon 1998; NRP, 2000) have noted the importance of teaching students comprehension strategies to enhance their comprehension. They recommend that teachers actually teach students how to comprehend material by using a combination of reading comprehension techniques. Some of these include: question answering, activating background knowledge, summarization, generating questions, comprehension monitoring, using graphic organizers, and teaching story structure. When students are able to use these strategies successfully, they perform better in recall, answering questions, generating questions, and summarizing texts. Perhaps had the experimenter taught the participants how to answer the various types of questions, their comprehension on the questions may have improved.

Research Question Three

*What effect will the RN computer software program have on the oral retell of the participants?*
After the participants completed the pass timing activity, the experimenter asked the participants to verbally state everything that he/she could remember about the story. A digital timer was set for one minute and the experimenter counted the total number of words that the participant could remember from the story. According to Good and Kaminski (2002), students need to meet the ORF benchmark goal for the second grade level and have a retell score of at least 25% of their ORF score in order to be on track for comprehension.

All six of the participants were able to increase the average number of words re-told from baseline to intervention. Five of the six participants doubled the number of words re-told during intervention; Quantasia was the exception. The average number of words re-told during baseline for all participants combined was 14. This number increased to an average of 32 during intervention.

Judging by the Good and Kaminski (2002) requirements, none of the participants are on track for comprehension, but this is due to their ORF levels on the RN stories. It is important to point out that Tamika (89 CWPM) was just shy of the 90 CWPM criteria for second graders. Considering that all of the participants were reading well below grade level prior to the intervention, their retell scores on the RN stories was still impressive.

Just examining the 25% criterion without the ORF component reveals that every participant except Angela would be on track for comprehension. Still, Angela would have only needed to retell an average of four more words per minute in order to be on track. Her needing to emit a higher level of responses is due to her average CWPM being 95 during treatment. Furthermore, Angela had a bit more difficulty than the rest of the participants due to her limited English proficiency in expressive language. She is much
better at receptive language as her comprehension and ORF scores show. Linan-Thompson and colleagues (2007) speculate that ELLs require more time to master newly acquired skills and need more opportunities to practice new vocabulary words and build expressive and receptive language skills. Gersten and Baker (2000) agree and propose that ELLs should be actively engaged in conversational and academic language in the classroom on a daily basis. Reading and language instruction for ELLs is a hot topic and surely will be the subject of much research to come.

Research Question Four

*What effect will the RN computer software program have on the oral reading fluency on generalization passages (i.e., AIMSweb) of the participants?*

*Correct Words per Minute*

These results were obtained by measuring how many words the participants read correctly during an unpracticed “cold” reading of a second grade level AIMSweb progress monitoring passage. The participants read the new story for one-minute and the experimenter collected data on how many words were read correctly. This number was calculated by subtracting the number of errors from the number of total words read.

The data reveals that every participant’s CWPM scores on AIMSweb probes increased from baseline to intervention, suggesting that this computer program was effective at increasing their ORF on new, unpracticed grade level passages. Collectively, the participants gained an average of 13 CWPM from baseline to intervention. Individual gains in CWPM correspond with the collective gains. Angela’s average individual gains at 23 CWPM were actually higher than the group average. Tamika also gained an average of 17 CWPM, which was a bit higher than the group average.
Even though all of the participants had high effect sizes (ES), comparing baseline and intervention scores, their actual average number of words read on the AIMSweb generalization probes were still lower than the 90 CWPM that second graders should be reading by the end of the school year. The two students who were approaching that goal were Angela and Tamika, with 65 and 67 CWPM, respectively.

One reason that the participants in this study may not have performed as well on these generalization passages is because it was the first and only time they encountered the passage. In some recent studies (Daly et al., 1999; Daly et al., 2005), the researchers actually used RR techniques to teach the generalization passages. In the 1999 study, the researchers used high content overlap (HCO) passages as one of the treatment components with RR or LPP. These HCO passages contained a considerable number of words from the instructional passages. As described in Chapter 2, the participants were provided one of the brief treatments on instructional level passages, assessed on the number of CWPM, and then assessed on a HCO passage. If there were no increases in CWPM on the HCO passages after receiving the intervention on the instructional passages in the RR/SM condition, the researchers then instructed on the HCO passages. The researchers expected the participants to perform well on the HCO passages due to the overlapping words; however, they found that three of the four participants needed instruction on these generalization passages in order to increase the number of CWPM. Therefore, the researchers stated that one cannot assume that ORF behavior will be generalized.

Their assertion seems correct in that there were instances during the intervention when the participants encountered the same words in the RN program and on AIMSweb
passages. One particular word was speed. The RN stories “Fun on Ice” and “Gliding Over the Snow” contained the word speed a total of three times. The participants heard the words during at least three read alongs with the computer, and read the word during all of their practice reading activities. Therefore, they heard and read the word speed a minimum of 11 times. Yet, in the AIMSweb generalization story “All Summer,” none of the participants read the word speed correctly. Thus, teachers must not hope that reading skills generalize, they need to train or teach them (Alber-Morgan, 2010).

**Incorrect Words per Minute**

Results on the number of IWPM were gathered by counting any errors (i.e., substitutions, omissions, additions, etc.) during the one-minute cold read and writing that number on the passage and data collection sheet. Only three of the participants decreased the number of errors made on AIMSweb passages from baseline to intervention, one participant’s errors remained the same (Quantasia), and two participants’ errors increased (Diana and Javon).

This does not come as a surprise as much of the research discussed herein (Begeny et al., 2006; Daly et al., 1999; Nelson et al., 2004) provides evidence for decreases in errors only when including an error correction procedure. Therefore, a stable range of errors was to be expected as error correction was not a part of the procedures for assessing generalization in this study. Welsch (2007) also found a steady range of IWPM on generalization passages for the students in his study. Even so, the students in the current study and the Welsch study all made improvements on CWPM on generalization passages. Despite these improvements, the failure to reduce error rates highlights the need for strategies that simultaneously address errors and speed.
Criterion Changes

As noted previously, the experimenter changed each participant’s goal criterion throughout the study by averaging the two highest probes from the previous criterion. The participants had to meet or exceed their goal for at least two RN probes in order to change their criterion. This was done to determine how increasing the number of CWPM would affect generalization scores.

Even though every participant made ORF gains from baseline to treatment on AIMSweb passages, unfortunately, there was not clear evidence as to how each criterion change affected these scores. However, what is clear is that more effects can be seen for Angela and Tamika, emerged as the two strongest readers in the study, with criteria over 100 CWPM, and CWPM treatment averages near or above the second grade benchmark of 90 CWPM.

Recent researchers (Gibson, 2009; Welsch, 2007) have begun examining the effects of changing goal criteria on ORF performance. Only 3 of the 4 participants in the Welsch study made limited ORF progress on generalization passages, so he decided to change their criterion by averaging the number of CWPM in the previous four generalization passages and increasing this amount by ten percent. After the students met the criterion for three out of four sessions, he increased their criterion by 10% more. In so doing, he noticed an increase on generalization passages in CWPM for all four students.

Gibson (2009) used a different changing criterion tactic. The participants in his study made progress in ORF on the treatment probes during the intervention; however, they were not making much progress on AIMSweb generalization passages. Therefore, he increased each of his participants’ goals by averaging the two highest probes from the
first phase and made that the goal for the second phase of the study. By increasing the
criterion, half of his participants’ CWPM on AIMSweb generalization passages increased
to over the DIBELS benchmark goal of 40 CWPM for first graders. Two more
participants read 30 or more CWPM, and all of the participants had large effect sizes for
CWPM on these passages.

Both the above-cited research studies and the current investigation found that
changing the participants’ goals by increasing the number of CWPM on treatment
passages also promoted gains in ORF on generalization passages. More research is
needed to determine the exact variables that promote greater fluency. Specifically, is
there a certain magnitude or length of the criterion change that is more or less effective at
promoting generalization?

Research Question Five

What effect will the RN computer software program have on the number of
comprehension questions answered correctly on generalization passages (i.e., AIMSweb)
by the participants?

AIMSweb progress monitoring probes do not have corresponding comprehension
questions, therefore, the experimenter constructed a five question quiz for each
AIMSweb passage. The quiz contained four multiple-choice questions and one open-
ended question, and followed the same format as the questions in the RN program. For
example, every question number three in the RN program is a vocabulary question and
says, “What does the word ______ mean in this story?” Of course, there was a
vocabulary word from the story in the blank (a full explanation of the type of
comprehension questions was presented in Chapter 3). Responses to these five questions on the quiz were used for the comprehension question data for AIMSweb passages.

The results for the percentage of comprehension questions answered correctly on AIMSweb passages were quite modest. Three participants’ scores increased from baseline to intervention, one decreased (Diana), and two remained the same. Again, it is important to note that each participant’s performance on this comprehension measure was commensurate with their performance in ORF on AIMSweb passages. For example, Diana scored quite low in ORF and regressed in comprehension; whereas Angela and Tamika scored quite high in ORF and were both the highest performers on AIMSweb comprehension, answering 75% and 83% correct, respectively.

As noted previously, there are a few reasons why the participants did not perform as well in answering comprehension questions as opposed to retelling the story. There were still a limited number of responses that the participants could make, the participants had never encountered many of the words they were reading in these generalization passages, and the experimenter did not directly teach the participants strategies on how to choose the correct answer.

Be that as it may, Markell and Deno (1997) believe that students need to read 90 CWPM in order to answer most (70%) comprehension questions from a story. In the current investigation, one-third of the participants were able to answer most of the comprehension questions from the AIMSweb generalization passages, even though their ORF was only approaching the 90 CWPM goal. This does not negate the importance of fluency as a precursor to comprehension by any means. It is simply an observation of the performance of two students who actually excelled on comprehension questions on
AIMSweb generalization passages. Future research in this area might clarify this assumption.

Research Question Six

*What effect will the RN computer software program have on the oral retell on generalization passages (i.e., AIMSweb) by the participants?*

After the participants finished reading the cold AIMSweb passage, the experimenter asked the participants to verbally state everything that he/she could remember about the story. A digital timer was set for one minute and the experimenter counted the total number of words that the participant could remember from the story.

All six of the participants were able to increase the average number of words re-told from baseline to intervention. Two participants, Diana and Angela, tripled the number of words re-told during intervention. The average number of words re-told during baseline for all participants combined was 14. This number increased to an average of 25 during intervention. Again, even though the participants were not reading 90 CWPM on these passages, the retell data themselves show that every participant is on track for comprehension on these cold AIMSweb generalization passages.

This is interesting because this was the first time that the participants encountered the passage, yet they were all able to remember much of what they read. One reason for this might be because they were not as focused on reading fast to beat their goal. While this is only speculation, other researchers (Gibson, 2009; Welsch, 2007) have also found that students’ retell fluency increased on generalization passages. Both researchers used changing criterion tactics in their studies as well. Word retell is definitely an area that needs to be researched further.
Research Question Seven

*What effect will the RN computer software program have on the oral reading fluency on generalization passages conducted in the classroom using the classroom curriculum?*

**Correct Words per Minute**

These results were obtained by measuring how many words the participants read correctly during an unpracticed “cold” reading of a passage from their DI classroom curriculum. The participants read the new story for one-minute and the experimenter collected data on how many words were read correctly. This number was calculated by subtracting the number of errors from the number of total words read. These probes were given at a reading table in the back of the classroom, while the teacher instructed the rest of the class on a different subject.

The data reveals that five of the six participant’s scores on CWPM on DI probes increased from baseline to intervention. Collectively, the participants gained an average of 12 CWPM from baseline to intervention. Javon and Angela’s average individual gains were actually higher than the group average, at 23 and 21 CWPM, respectively. Tamika actually scored three points lower in CWPM during treatment, but she also had the highest ORF level of all six participants at the beginning of the study, and she did not receive as much intervention as the other participants because she was in tier three. Of all of the literature reviewed, only one study assessed generalization in the classroom. Hitchcock and colleagues (2007) collected ORF and comprehension data in the classroom at one and six months after the intervention. Their results revealed that both fluency and comprehension gains were maintained. What is lacking from their study is whether or not the assessment measure was from the intervention material or the
classroom material. This is an area of much needed research, as generalization is not a well-studied topic. Generalization to classroom materials is of particular importance because interventions that are provided in the clinical setting may enhance ORF on those measures, but the ultimate goal would be for those gains to be realized in the classroom as well. Further research is warranted.

Incorrect Words per Minute

Results on the number of IWPM were gathered by counting any errors (i.e., substitutions, omissions, additions, etc.) during the one-minute cold read and writing that number on the passage and data collection sheet. Only three of the participants decreased the number of errors made on DI probes from baseline to intervention and the other three participants’ errors remained the same. These results parallel the research presented on IWPM in question four.

Criterion Changes

Originally, the experimenter was going to probe the DI classroom material every other criterion change, however, this would not have been enough probes. Therefore, once the experimenter realized that it was going to be difficult to gather a representative number of probes in that manner, she administered one DI probe for every criterion change from the second criterion change until the end of the study.

Even though most of the participants made ORF gains from baseline to treatment on the DI probes, there was not clear evidence of how each criterion change affected the DI scores. The participants’ scores in ORF for DI passages generally corresponded to their individual performance in ORF on AIMSweb and RN passages. More research is
needed in order to determine how changing the criterion in the clinical setting impacts the ORF in the classroom or other generalization settings (i.e., home, in the library).

Research Question Eight

*What effect will the RN computer software program have on the number of comprehension questions answered correctly on generalization passages conducted in the classroom using the classroom curriculum?*

The DI curriculum did have corresponding comprehension questions for each story; however, not all of the comprehension questions followed the same format as in the RN program. Therefore, the experimenter constructed a five question quiz for each DI passage. The quiz contained four multiple-choice questions and one open-ended question, and followed the same format as the questions in the RN program. Oftentimes, the experimenter was able to use two to three questions directly from the DI program (e.g., question one asking about the main idea of the story). Responses to these five questions on the quiz were used for the comprehension question data for the DI passages.

The results for the percentage of comprehension questions answered correctly on the DI passages were also quite modest. Three participants’ scores increased from baseline to intervention, and three participants’ scores decreased. A major limitation to point out is that students in tiers one and two only had one DI probe during baseline, whereas tier three students actually had two DI probes during baseline. This difference might have had an enormous impact on Javon’s scores decreasing from 80% correct during baseline to only 35% correct during intervention. That particular baseline probe may have been an interesting story or quite easy for Javon, which would magnify his
comprehension. However, Javon’s scores during the intervention appear to be more similar to his typical responding.

Nevertheless, it is surprising that the participants did not score higher on the DI stories because the curriculum often presents a number of passages in a row that has the same story line and characters in a drawn-out plot. Because the participants might have already been familiar with the storyline, one might believe that their understanding of the story would increase and thus, show improved comprehension scores. These findings were only partially supported as just half of the participants’ scores increased on the DI classroom curriculum passages.

Once more, there was only one study that included a generalization component in the classroom (Hitchcock et al., 2007). Nonetheless, their results showed that comprehension gains were maintained up to six months after the study. However, the researchers did not discuss exactly what measure they used to assess comprehension, thus, limiting comparisons between the current investigation and their study. Much more research in this area is needed.

**Research Question Nine**

*What effect will the RN computer software program have on the oral retell on generalization passages conducted in the classroom using the classroom curriculum?*

After the participants finished reading the cold DI passage, the experimenter asked the participants to verbally state everything that he/she could remember about the story. A digital timer was set for one minute and the experimenter counted the total number of words that the participant could remember from the story.
All six of the participants were able to increase the average number of word re-told from baseline to intervention. Diana quadrupled and Tabitha doubled the number of words re-told during intervention. The average number of words re-told during baseline for all participants combined was 19. This number increased to an average of 33 during intervention. Individual averages reveal proportionate responding. Yet again, these data show that every participant is on track for comprehension on these cold DI passages.

These are likely results considering that many of these words had been encountered before and many of the words in the stories were decodable. The classroom curriculum (DI) uses fiction passages that are written in a predictable manner and have a lot of words that are repeated for meaningful practice. As a result, the participants might find it easier to recall many of the words in the story. Unfortunately, there is no literature to support this conclusion, as the research-base for word retell is non-existent. Researchers might find it useful to investigate word retell fluency more fully as a measure of reading comprehension.

Research Question Ten

To what extent will the participants respond favorably to the RN program?

The students answered questions about their motivation, enjoyment of the computer program, and their reading abilities. All six of the participants responded “yes” to each question, meaning that they liked working on the computer and felt they became better readers. This is not surprising because five of the participants were sad that the intervention was over and that they could not work on the computer anymore. Additionally, the participants were sad when it was not their day for the intervention and would ask the experimenter when they would get to come next. The fact that the
participants received individualized instruction on a novel computer program with a person other than their teachers could have played a role in their excitement to participate in the intervention.

When Diana was asked why she liked reading on the computer, she responded, “Because it’s fun to do and when you grow up you can be a teacher and read to children. If you play around, you won’t be able to read.” This response was quite interesting because she was by far, the lowest reader in the study and made the least progress. Yet even she realized the benefit of reading interventions.

Research Question Eleven

*How will the teacher rate the effectiveness of the RN intervention on the ORF and comprehension of the students?*

Two teachers rated the effectiveness of the program by answering a questionnaire with five questions pertaining to the intervention. Questions one and two asked their opinion on their students’ reading achievement by the end of the intervention. Questions three and five intended to assess the teachers’ opinion on the importance of reading interventions for struggling readers, and question four assessed their students’ enjoyment of the program. There was space at the bottom of the questionnaire for comments.

Teacher 2, the participants’ second-grade teacher, noted that the experimenter did a great job with the kids. She thought the program was helpful and that the students could not wait until it was their turn to get on the computer. She did believe, however, that some students put forth more effort than others, and that it showed. She commented that the students were trying harder in class and that their confidence had improved. She believed that had the students had even more time to continue with the program, then
they would have become even better readers. Teacher 1, the reading teacher for three of the participants, commented that her students’ overall reading has extremely improved. Both teachers agreed that intervention programs are very important for struggling readers and they would allow their students to participate in similar intervention programs in the future.

The social validity information collected from the students and teachers are in line with research from Hasbrouck et al. (1999). The teachers in their study pointed out that the students who used the RN program were more confident, had higher self-esteem and were hopeful about their future reading success. Additionally, the students valued their performance feedback in the form of graphs. This research supports the use of scientifically-validated research techniques that promote ORF and other intrinsic factors (i.e., self-esteem, confidence).

Discussion of Secondary Dependent Variables

The WJ-III and DIBELS were used as secondary measures to determine the effects of the intervention on the participants’ overall reading ability. The WJ-III is a standardized measure that compares each participant to a normative sample of his/her grade level peers. The participants’ raw scores on pre- and post-tests were used to determine their grade level equivalents, standard scores, and classifications. The DIBELS benchmark assessment scores were used to determine the participants’ risk status before and after the intervention. See Appendix L for classification and risk status information.

Four sub-tests of the WJ-III were used in this study: LWID, WA, RF, and PC. Results for the LWID sub-tests revealed that all of the participants were classified in the low to average range at pre- and post-test. Five of the participants’ grade equivalent
scores increased, and one remained the same. Angela obtained the most dramatic increase in scores, from 1.8 to 2.7, which classified her in the solid average range.

Five of the six participants increased their scores on the WA sub-test, with Tamika being the only participant to regress from pre- to post-test, from 2.9 to 2.7, respectively. All of the participants’ post-test scores placed them in the low-average to average range.

Results on the reading fluency sub-test were more dramatic for some participants than others. Javon increased his scores from <K.0 at pre-test to 1.9 at post-test, with a low to low-average classification. Angela also increased from 1.5 to 2.5 by the end of the study, which places her in the average range. Three other participants also increased grade equivalent scores from pre- to post-test, whereas Diana’s scores actually decreased from 1.8 to 1.4.

All six participants were able to increase their grade equivalent scores from pre-to post-test on the PC subtest. However, the increases in scores still resulted in classifications from very low to average.

Even though most participants made gains from pre- to post-test on the WJ-III, these results must be interpreted with caution. Standardized measures tend not to be as reliable for younger children, such as the participants in the present study. Furthermore, it is quite difficult to change standardized scores in such a short amount of time. Last, as the test takes the students’ age into consideration when scoring, participants could have increased their performance, but this progress would not be shown at the end of the study because the scoring scale had increased based on the students’ current age.
Results on the DIBELS benchmark assessments revealed that all six participants were able to increase the number of CWPM from pre- to post-test. However, only one participant, Angela, was able to reduce her risk status from “at risk” to “some risk.” She also made the most improvement, with an increase of 51 CWPM from pre- to post-test. Four other participants also made double-digit gains in CWPM, but Diana only gained 4 CWPM. Even though Angela was the only participant to reduce her risk status, four other participants made substantial gains as well. Again, the results appear promising for most of the participants, however, the need for further reading interventions is evident.

Limitations

Despite the fact that the participants in the current investigation improved their ORF and, in general, their comprehension, there were some limitations to the study. Therefore, some caution is warranted when interpreting these findings.

Intervention Implemented in the Clinical Setting

The results of this study support the use of RN as an effective intervention for students who are at-risk for reading failure. However, the RN intervention was implemented in an empty classroom with only the experimenter and another student. Even though there was noise in the hallway and an adjacent classroom, the room was virtually free from distractions. The experimenter was also available for the participants at any time throughout the intervention.

These conditions are not typical in most classrooms, especially in urban schools. There is a minimum amount of supervision that is needed in order to ensure that the students interact appropriately with the program. The effectiveness of this program may differ if implemented in the natural classroom environment under less supervision; or
even in a computer lab when the whole class might be using the RN program at the same time. Future research is needed to determine the effectiveness of RN under more typical classroom conditions.

*Comprehension Questions for Generalization Passages*

There were two measures of comprehension in the current study. One was the number of comprehension questions answered correctly and the other was the number of words re-told per minute. Having two measures of comprehension provides a more well-rounded view of comprehension. However, the experimenter developed all of the comprehension questions for AIMSweb generalization passages and some of the questions for DI generalization passages.

Even though all attempts were made to develop questions that corresponded with the readability of the passages, it is quite possible that the AIMSweb questions were more difficult for the participants. In addition, the content of the DI stories was so limited that the experimenter often found it difficult to develop questions that fit the RN format. Therefore, there was probably more variability in the readability of the generalization passage comprehension questions versus the RN program.

*Generalization in Classroom Reading Groups*

The data for the DI generalization probes revealed that five of the six participants’ scores increased from baseline to treatment. These results are consistent with teacher reports of improved reading performance in the classroom. The experimenter administered the DI generalization passages in the participants’ actual classroom. However, it was not during their daily reading block and the teacher did not administer the probes. Ideally, generalization in the participants’ classroom should be assessed by
their classroom teacher under normal classroom conditions. The experimenter tried to simulate “normal classroom conditions” by sitting in the back of the classroom during other teacher instruction. Future researchers might utilize actual classroom teachers to assess generalization of reading performance.

**Limited Intervention Time**

As with any research conducted in school settings, there are usually issues with logistics and scheduling. The intervention was brief considering that there were missed sessions due to calamity and professional development days; Spring break; experimenter illness; student absences; and limited computer availability at the beginning of the study. In addition, the school only provided the experimenter with a one-hour time slot per day in which to implement the intervention. The RN stories took the participants anywhere from 20-45 minutes to complete the whole teaching sequence, therefore, no more than two to four students could receive the intervention per day. Consequently, the experimenter and Teacher 2 suspect that had the intervention been implemented for a longer period of time, for instance, six months of the school year, then the participants would have made even more progress on ORF and comprehension. Future studies of longer duration would be able to support this claim.

**Implications**

Results from this study provide more support for the effectiveness of the RN software edition. Limited research has been conducted with RN (Denton et al., 2004; Denton et al., 2006; Hasbrouck et al., 1999), and with the software edition in particular (Gibson, 2009). The current investigation extended the research by Gibson, and supports
the results obtained by Denton and colleagues (2006) and Hasbrouck and colleagues. In spite of this, there are some issues that might impact implementation.

One issue is with the level of student independence during implementation. All of the participants knew how to use a computer prior to the study and only one student (Angela) needed two sessions to learn how to use the RN program. The participants were all 100% capable of progressing through the program alone; however, on various occasions, the experimenter needed to instruct the participant to click on a word that he/she thought they had read correctly, but had not. This is of concern because struggling readers might start to practice errors if they do not realize that they are reading the word wrong. Therefore, it is important that teachers are available to monitor the students in order to provide needed prompts while using the RN program.

Teachers would also need to be available to monitor the program to ensure that students do not “cheat” to try to beat their goal. On one occasion, the experimenter observed Diana, the lowest performing student in the study, trying to manipulate the computer so that she obtained a higher than actual performance. When the bell sounds at the end of one minute, the computer instructs the participant to click on the last word read in order to calculate the number of CWPM. Diana tried to click further in the story than she had actually read. However, she clicked too many words ahead, so the experimenter questioned her as to how she read almost 20 more words per minute during that practice read as opposed to the others. Diana admitted what she did and the experimenter had a discussion with her about how that type of behavior only hurts her because it is not helping her learn how to read the words that she did not know. The participant understood and did not engage in that type of behavior again. This highlights the need for
teacher supervision, to ensure that the participants implement the RN program with fidelity so that they can receive the best benefit possible.

Additionally, teachers must be present during the implementation of the RN program because it is a computer program, and technology is not without its flaws. However, there were not many technological issues over the course of the study. On three occasions the program shut down while the participants were reading, but when the experimenter restarted the program, it began in the same place where they left off. On two occasions the computer did not present the “next” button after the third read along; and on one occasion the computer did not register the first practice reading. Considering that the participants interacted with the RN program over 100 times throughout the course of the study, these were minimal technological glitches. Still, these minor technological problems highlight the need for teacher supervision of this intervention program.

A feature of the RN program that is quite beneficial for classroom teachers is the ability to individually adjust goals and reading levels. Today’s classrooms are filled with students of varying reading ability levels. Teachers can use this RN program to provide individualized reading practice for more students at once. It is quick and simple to adjust the goal (CWPM) by student, so teachers do not have an additional time burden by providing this individualized instruction. Furthermore, the students can work at their own pace (Lonigan et al., 2003; McCullough, 1995) with interesting and motivating activities (Macaruso et al., 2006), and are provided with many opportunities to repeatedly practice skills to build fluency (Macaruso et al.) and mastery in an environment that is not threatening or embarrassing (Diem & Katims, 2002; McCullough). Those are likely some
of the reasons why the participants in the current investigation became more confident in their reading abilities.

Yet another benefit of this program is the repeated reading aspect in the form of the read along and practice reading activities. Previous discussion has already testified to the importance of RR in promoting ORF in struggling readers, however, this type of practice might also prove extremely beneficial for ELLs. The ELL in this study, Angela, actually had the highest ORF scores on RN probes, DI generalization probes, the DIBELS ORF benchmark assessment; and the second highest ORF scores on AIMSweb generalization passages and on the WJ-III ORF subtest. She was the only participant to reduce their risk status according to DIBELS. It is quite possible the constant hearing and seeing of words in English provided Angela with the tools she needed to start excelling in reading. This is not far-fetched as Linan-Thompson et al. (2003) reported that 66% of the ELL students in their study progressed rapidly through their reading intervention lessons; learning, mastering, and generalizing their newly acquired reading skills after minimal practice. The researchers assert that this provides evidence that the ELLs just had not received good instruction in English as opposed to the students having a reading disability. There is limited generality of the results of the current investigation as there was only one participant who was an English language learner. More research needs to be conducted with the RN program and ELLs, providing more evidence of its effectiveness with populations other than native English speakers.

Future Research

The current investigation of RN (software edition) presents promising results in the use of CAI in reading for elementary school students who are at-risk for reading
difficulties. Even though the effects of the RN program produced robust effect sizes for ORF and even gains in comprehension for some students on treatment and generalization passages, there is still a need to further investigate this area of research.

First, future research should be conducted with the RN software edition. To date, there are limited studies regarding the effectiveness of the RN program (Denton et al., 2004; Denton et al., 2006; Hasbrouck et al., 1999), and more specifically, the software edition (Gibson, 2009). Although the RN program uses scientifically-validated teaching techniques, further research must explore the effects of these techniques on the reading skills of early primary aged students, as well as ELLs. Furthermore, can the effects of the RN program be generalized into other settings, other than the clinical setting; and other materials (i.e., classroom curricula, AIMSweb passages)? While the results of this intervention show promise, more research is warranted.

Certainly, more research needs to be conducted on the effects of changing the goal criterion to promote generalization of reading fluency. Results from the current investigation appear promising that by systematically and individually increasing the CWPM goal on treatment passages in the RN program; there was an increase in unpracticed generalization passages. The question of “how” remains. Specifically, is there a certain magnitude of change that appears to be most beneficial to promote generalization? Surely, the length of the phase of each criterion change might also play an important role in the ability to generalize. The scientific community could benefit from learning more about these systematic increases and the effects on generalization.

Some questions also remain about this program’s effectiveness on reading comprehension. Only half of the students in the present study made very modest gains in
comprehension scores on the RN probes. Analysis of comprehension scores on both measures of generalization was quite similar. Future researchers might study the link between ORF and comprehension. In particular, is there a certain number of words that the participants must read in this program in order to master the comprehension component? Some researchers (Ardoin et al., 2008; Bonfiglio, 2004; Markell & Deno, 1997) have discussed a threshold for reading fluency to impact comprehension and generalization. More research is needed to confirm these findings, especially in regards to the RN program.

Last, this study was very brief. Teacher social validity data revealed the belief that even more ORF gains could have been made had the intervention lasted for a longer period of time. To be sure, questions remain about the length of intervention time needed in order to affect the most substantial and lasting amount of change. Additionally, could comprehension have been a lagging factor in that more increases might have been documented over a longer period of time? That is definitely an empirical question that begs to be answered.

Summary

This study examined the effectiveness of the Read Naturally (Software Edition) on the ORF and comprehension of second graders who were at-risk in reading. Six students from an urban elementary charter school participated in the study. They were included based on teacher recommendations of needing intervention for reading fluency, and DIBELS scores confirmed this as all six students were “at-risk” on the winter benchmark assessment. The experimenter used a multiple-baseline design with changing criterion tactics to determine the effects of the RN program on the participants’ ORF and
comprehension on the RN program and two measures of generalization. Standardized instruments and curriculum-based assessment measures were collected as pre- and post-test measures in order to provide more information on the participants’ reading growth.

Results from this study were mixed. All six participants made gains in ORF on the RN and AIMSweb generalization probes. Five of the six participants increased their ORF on the DI generalization probes in the classroom. This resulted in large effect sizes for most of the measures for most of the participants. Two measures of comprehension were obtained: percentage of comprehension questions answered correctly and the number of words re-told per minute. Scores on the comprehension questions were not as robust as scores on ORF and word retell. All participants increased the number of word re-told per minute across all three types of probes (RN, AIMSweb, DI), but on any given type of probe, only three students increased the number of comprehension questions answered correctly. Although all participants made gains from pre- to post-test on DIBELS, only one participant was able to reduce her risk status. Furthermore, even though most participants made pre- to post-test gains on the subtests of the WJ-III, these increases still classified them in the low to average range.

In spite of these mixed results, this study demonstrates the effectiveness of the RN program as a supplemental intervention for students who are having difficulty in reading. The students in this study were already reading well below grade level and five of the six of them made substantial gains in ORF. This is encouraging because this brief intervention provided a boost in performance that helped to start to narrow the gap between these low-performing students and their grade level peers. Moreover, this intervention was provided through CAI, which offers teachers greater flexibility in
providing individualized reading instruction to more learners at the same time. In this era of RtI, CAI such as RN could be beneficial to implement with learners in any tier. Thus, CAI in reading might be part of the answer to reducing reading risk for students in classrooms across the country.
REFERENCES


system on first-grade students’ reading gains. Reading and Research Quarterly, 21, 361-376.


Coyne, M. D., Kame’enui, E. J., Simmons, D. C., & Harn, B. A. (2004). Beginning reading intervention as inoculation or insulin: First-grade reading performance of


Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and
instruction: What, why, and how? The Reading Teacher, 58(8), 702-714.


Lalas, J. (2007). Teaching for social justice in multicultural urban schools:


language learners with reading difficulties. *The Elementary School Journal, 103*(3), 221-238.


Martens, B., Eckert, T., Beeny, J., Lewandowski, L., DiGennaro, F., Monarello, S.,
on the reading performance of low-achieving second and third grade students.
*Journal of Behavioral Education, 16*, (1) 38-53.

fluency. *Intervention in School and Clinic, 34*(5), 278-283.

strategies for first-grade readers with and without additional computer-assisted

McCardle, P. & Chhabra, V. (Eds.). (2004). *The Voice of Evidence in Reading Research*

McCullough, C. (1995). Using computer technology to monitor student progress and

intervention on literacy development in multiple-language settings. *Learning
Disabilities Quarterly, 30*(3), 197-212.

Miller, J. & Schwanenflugel, P. J. (2008). A longitudinal study of the development of
reading prosody as a dimension of oral reading fluency in early elementary school

Examination of the predictive validity of preschool early literacy skills. *School


Skiba, R. J., Poloni-Staudinger, L., Simmons, A. B., Feggins-Azziz, R., & Chung, C. G.


Digest, 62(6), 47-49.


APPENDIX A

Example of a RN story and comprehension questions
Story 3 – Amelia Earhart

Amelia Earhart was a little girl in the early 1900s. In those days, girls wore only dresses. They were told to walk slowly. They were told to sit still. Amelia was not like other girls. She wore pants. She played ball and climbed trees. Amelia even wanted to fly airplanes. So she went to flight school to learn to fly.

In 1928, she flew across the Atlantic Ocean. She was the first woman to make that flight. In 1932, she made the same flight. This time she went alone. She was the first woman to do these things.

In 1937, she tried to fly all the way around the world. Her plane disappeared in the Pacific Ocean. Amelia and her plane were never found. Amelia was a brave person.

F. K. Grade Level 3.8
Amelia Earhart Comprehension Questions

1. What is most of this story about?
   A. girls in the early 1900s  
   B. a brave woman  
   C. learning to fly  
   D. an airplane that disappeared at sea  

2. What did Amelia do before any other woman?
   A. flew around the world  
   B. wore pants and played ball  
   C. went to flight school  
   D. flew across the Atlantic  

3. What does the word flight mean in this story?
   A. an airplane trip  
   B. a set of stairs or steps  
   C. a run from danger  
   D. a trip around the world  

4. How was Amelia different from other girls?
   A. She walked slowly.  
   B. She sat very still.  
   C. She went to school.  
   D. She wanted to fly airplanes.  

5. What things did Amelia Earhart do that showed she was brave?  

3.6
APPENDIX B

Example of a DI story and comprehension questions
The Magic Pouch

The elf told the little girl, “When you are bad, the pouch will be bad to you.”

The girl picked up the pouch. She said to the elf, “I have been good to you. Let’s see if this magic pouch will be good to me.”

She reached inside the pouch and found ten round rocks that shine. “These round rocks are gold,” she shouted. “I’m rich.”

So the girl thanked the elf for the pouch.

Then the girl and her hound started down the tall mountain. They went down and down. They went into the clouds. When they left the clouds, the girl could see the ground. Down and down they went.

When they reached the bottom of the mountain, the sun was setting. It was getting late. The girl was tired. But she ran to her house.
Her mother met her at the door. She said, “Where were you? Your father and I have looked all around for you.”

The little girl did not tell her mother where she went. She said, “I went to sleep in the grass. I just woke up.” She told a lie, and that was bad.

F.K. Grade Level 0.3
Pouch Comprehension Questions

1. What is most of this story about?
   A. An elf playing a game.
   B. The girl and a magic pouch.
   C. The girl and her hound.
   D. An elf that lies.

2. What was in the magic pouch?
   A. An elf.
   B. 10 pieces of candy.
   C. 10 round rocks of gold.
   D. A doll.

3. What does the word pouch mean in this story?
   A. a small bag
   B. to poke your lips out
   C. a place to carry a baby kangaroo
   D. to swallow something

4. Why did the girl say that she was rich?
   A. Because the elf told her she was rich.
   B. Because her mom gave her money.
   C. Because she found gold buried in the ground.
   D. Because there were rocks of gold in the pouch.

5. Why did the girl's mom ask her where she was?
APPENDIX C

Example of AIMSweb story and comprehension questions
All summer I’ve wanted the three-speed mountain bike in the bike shop window. It is shiny black. It has hand brakes. It even has a holder for my water bottle. I spent every Saturday this summer admiring the shiny, black, three-speed bike.

Today Dad drove a friend and me to the bike shop. In my pocket, I had the money I saved for the bike.

“I will buy the shiny, black, three-speed bike today!” I said to my friend.

“Maybe you should look around first,” Dad said.

When I entered the bike shop, I saw many other bikes. I told the shop owner that I wanted to buy the shiny, black, three-speed bike.

Maybe you should look around first,” the owner said.

I looked at a bright green dirt bike. It had cool tires and a sticker with the word “Fearless.” I looked at a big red bike. It was a ten-speed bike that was very light. The bike shop owner said it could travel very fast. I looked at scooters and trikes. I looked at bikes with training wheels and bikes with only one wheel.

Finally, it was time to buy a bike.

“Do you know which bike you want to buy?” Dad said.

I looked at the bikes. I looked at my friend. I looked at the bike shop owner. “I want the shiny, black, three-speed bike!” I said.
All Summer Comprehension Questions (0.0)

1. What is most of this story about?
   A. A boy and his friend.
   B. How to ride a bike.
   C. A boy who wants a bike.
   D. When to go shopping for a bike.

2. What kind of bike did the boy want?
   A. A dirt bike.
   B. A 3-speed mountain bike.
   C. A red bike.
   D. A scooter.

3. What does the word shiny mean in this story?
   A. smooth and bright
   B. colorless
   C. dull
   D. ugly

4. How did the boy have enough money to buy the bike?
   A. His dad paid for it.
   B. The boy saved his money.
   C. The shop owner gave it to him.
   D. The boy’s friend bought it.

5. Why did the dad tell his son to look around the shop before buying the bike?
APPENDIX D

Experimenter Data Collection Sheet
Data Collection Sheet

<table>
<thead>
<tr>
<th>Date</th>
<th>Passage #</th>
<th>CWPM</th>
<th>IWPM</th>
<th>Retell</th>
<th>Comp</th>
<th>Type of Passage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX / GEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX / GEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX / GEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX / GEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX / GEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX / GEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX / GEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX / GEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX / GEN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX / GEN</td>
</tr>
</tbody>
</table>

Key:
- Passage # - the passage being presented
- CWPM - the number of correct words read during the one minute timed reading
- IWPM - the number of incorrect words read during the one minute timed reading
- Comp - the number of comprehension questions answered correctly
- Retell – the number of words the participant retells from the passage
- TX – Treatment probe
- GEN – Generalization probe
APPENDIX E

Computer Training Procedural Integrity Checklist
### Procedural Integrity – Training Checklist

<table>
<thead>
<tr>
<th>Trainer’s Actions</th>
<th>Yes</th>
<th>No</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduced program</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asked/checked if participant knows how to use the mouse pad</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructed to click on story</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Introduced key words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Introduced cold timing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Introduced unknown words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prompted to click last word read</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prompted to click unknown words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explained goal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explained read along</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Explained practice reading</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asked if there were any questions about the program</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX F

Procedural Integrity – RN Computer Checklist
<table>
<thead>
<tr>
<th>RN Computer Actions</th>
<th>Yes</th>
<th>No</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presented selected story</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presented all key words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timer started when clicked</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Displayed correct goal and cold timing scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presented story three times during read along</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presented the “next” button after the third read along</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presented all comprehension questions during the quiz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scored incorrect questions as such</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prompted the participant to keep practicing until a teacher comes for the pass timing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presented password protection for the pass timing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correctly scored all measures</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX G

Procedural Integrity – Data Collection Sheet
Procedural Integrity Checklist – Data Collection

<table>
<thead>
<tr>
<th>Presented story</th>
<th>Yes</th>
<th>No</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructed participant to read as fast and/or best as he/she can</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Told participant would provide the correct word if it was unknown</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provided all unknown or mispronounced words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructed participant to remember everything about the story</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Started/stopped timer correctly</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX H

Procedural Integrity – Participant Use of RN Checklist
<table>
<thead>
<tr>
<th>Activity</th>
<th>Yes</th>
<th>No</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clicked key words</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read along three times</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attended during read along (i.e., oriented at screen, tracking)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clicked unknown words during practice reading</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clicked last word read after the bell sounded</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read the questions and all answers before clicking on an answer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asked for help, if needed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clicked on all vocabulary words during read along (1-3 times)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX I

Social Validity Questionnaire – Teachers
Social Validity Questionnaire – Teachers

1. Do you think that the student’s overall reading has improved since completing the intervention ________
   a) A lot
   b) Somewhat
   c) A little

2. To what degree would you say that the student became a more fluent reader after the intervention _____
   a) Very
   b) Somewhat
   c) Not very

3. Would you allow any of your students to participate in similar programs in the future?
   a) Yes
   b) No

4. Do you think the students enjoyed this reading intervention ______
   a) A lot
   b) Somewhat
   c) A little

5. How important do you think reading intervention programs are for struggling readers?
   a) Very
   b) Somewhat
   c) Not very

Comments:
APPENDIX J

Social Validity Questionnaire – Students
Social Validity Questionnaire – Students

1) Did you like reading on the computer?
   a. Yes
   b. No

   Why?

2) Do you think you became a better reader?
   a. Yes
   b. No

   Why?

3) Did you like working for prizes?
   a. Yes
   b. No

   Why?

4) Would you have worked on the computer even if you did not get prizes?
   a. Yes
   b. No

   Why?

5) Would you like to continue to read on the computer?
   a. Yes
   b. No

   Why?
APPENDIX K

Sample Read Naturally Lesson
Story Selection

Story Selection Screen

Each Read Naturally level includes 24 stories in two sets of 12. The set of available stories is based on the reading level the teacher chooses for the student.

Students work on the stories in a level until they pass all the stories, or they are ready for a new level. Allowing students to choose their own story provides them with a sense of ownership over their work.

Screen Highlights

1. Stories that the student has completed are marked "Passed." Students must complete a story before selecting a new one.
2. Students click a picture to select a story from the ones available at their reading level.
Key Words

Abraham Lincoln

Key Words
Civil War
lawyer
politics
self-made

1

lawyer

A lawyer is someone who works with the law.

John needed a lawyer because he broke the law.

2

3
The Key Words step provides students with vocabulary instruction, teaching them key words that are used in the story. They learn how to pronounce the words and what they mean.

Learning and saying these key words helps students read them with greater confidence when they encounter the words in the story. The Key Words step also encourages students to think about what they will be learning from the story.

1. Students listen and read along as the key words are read for them.
2. Students click each of the words to see and hear their definitions.
3. Students click the Next button when they are ready to move to the next step.
Peter Farrell, you read 70 words per minute.

Cold Timing Graph
In the Cold Timing step, the program times students for one minute as they read the story for the first time. This establishes a baseline for measuring students' improvement as they work on the story.

As students read, they should click words they don't know or have problems reading. Clicking on these difficult words increases students' awareness of unknown words and alerts teachers to words or word patterns that need to be taught.

1. Students can turn a reading guide on or off before they start reading. The guide highlights the line of text under the cursor, helping the students' eyes track across the line.
2. Students click on the stop watch button to begin the one minute timing.
3. Students read the story aloud and click words they don't know or have problems with. The words they click on change color and are subtracted from the final cold timing score. After the timing is complete, students can click the highlighted words to hear them spoken and practice reading them.
4. Students can click Finished if they finish the story before one minute has passed. If students click Finished too early (before they could have completed reading the story), the program may ask them to read again.
5. The Cold Timing graph shows the number of words the student read correctly, subtracting the number of difficult words the student clicked on from the total number of words read. The screen also displays the student's score and goal.
Abraham Lincoln was born in Kentucky in 1809. His mother died when he was 9 years old. Abraham was tall and strong. He worked hard and helped his neighbors. He loved to wrestle and tell stories. He only went to school for a year. Abe learned to read by himself.

When Abe grew up and left home, he had many jobs. He took a flatboat down the Mississippi River to New Orleans. He worked in a lumber mill. He was a storekeeper. Then he was a postmaster. At the time he kept learning. He read books about the law. Abe wanted to be a lawyer. In 1831, Abe passed a test and became a lawyer.

Then he got into politics. In 1860, Abraham Lincoln was elected the President of the United States. President Lincoln became one of our greatest presidents. He led the United States in the Civil War. He is remembered as the president who freed the slaves. And Abraham Lincoln was truly a self-made man. When he was killed in 1865, it was a great loss for America.
In the Read Along step, students read along while listening to a recording of the story, usually three times. (When setting up individual students’ story options, teachers can specify the number of read alongs, as well as the speed of each read along.)

This step helps students learn new words and encourages proper pronunciation, expression, and phrasing. Students should quietly subvocalize as they read to ensure that they actually read along.

1. When the read along begins, each sentence in the story is highlighted as it is being read. Students quietly read along with the recording.
2. After each read along, students can click the blue vocabulary words to see and hear their definitions.
3. After completing the required number of read alongs, students can either click Start to read along again for more practice, or click Next to continue to the next step.

Practice
In the Practice step, students are timed as they practice reading the story without the recording. They are required to practice until they can read the story at or above the goal rate that the teacher has set for them. It may take students from three to ten practice readings to reach their goals.

1. A box displays the number of words per minute the student read. Students who did not meet their goals must read again.
2. The teacher can review details about the student's progress through the story (for example, a list of difficult words the student clicked on in the Cold Timing and Practice steps).
3. The student can return to the Read Along step, continue practicing or, if the goal has been met, move to the next step.
Abraham Lincoln was born in Kentucky in 1809. His mother died when he was 9 years old. Abraham was tall and strong. He worked hard and helped his neighbors. He loved to wrestle and tell stories. He only went to school for a year. Abe learned to read by himself.

When Abe grew up and left home, he had many jobs. He took a flatboat down the Mississippi River to New Orleans. He worked in a flour mill. He was a storekeeper. Then he was a postmaster. At the time he kept learning. He read books about the law. Abe wanted to be a lawyer. In 1859, Abe passed the bar and became a lawyer.

Then he got into politics. In 1860, Abraham Lincoln was elected the President of the United States. President Lincoln became one of our greatest presidents. He led the United States during the Civil War. He is remembered as the president who freed the slaves. And Abraham Lincoln was truly a self-made man. When he was killed in 1865, it was a great loss for America.

1. What is the main idea of this story?
   A. Lincoln was a self-made man who became a great president
   B. Lincoln was the President during the Civil War.
   C. Abraham Lincoln was truly a self-made man. When he was killed in 1865, it was a great loss for America.

2. Why do we say Abraham Lincoln was a self-made man?
   He became successful with not much help from other people.
In the Quiz step, students answer up to nine comprehension questions about the story. The program checks the answers to the multiple-choice questions and repeats questions that the student answers incorrectly. During the Pass step, the teacher reviews and corrects the open-ended questions.

1. Students can review the story text and click on vocabulary words prior to answering the questions.
2. Students answer each of the questions in order. Each question focuses on a specific comprehension skill. For example, Question 1 tests whether the student understood the main story idea. The Comprehension Report enables teachers to track a pattern of errors and provide additional instruction in those areas.
3. Students type answers to open-ended questions using their own words. Teachers review and correct these questions during the Pass step, using the *SE Answer Guide.*
Abraham Lincoln was born in Kentucky in 1809. His mother died when he was 9 years old. Abraham was tall and strong. He worked hard and helped his neighbors. He loved to wrestle and tell stories. He only went to school for a year. Abe learned to read by himself.

When Abe grew up and left home, he had many jobs. He took a flatboat down the Mississippi River to New Orleans. He worked in a lumber mill. He was a storekeeper. Then he was a postmaster. At the time he kept learning. He read books about the law. Abe wanted to be a lawyer.

In 1836, Abe passed a test and became a lawyer. Then he got into politics. In 1860, Abraham Lincoln was elected the President of the United States. President Lincoln became one of our greatest presidents. He led the United States in the Civil War. He is remembered as the president who freed the slaves. And Abraham Lincoln was truly a sad event. When he was killed in 1865, it was a great loss for America.
At the Pass step, teachers listen to students read and determine whether they are ready to move on to the next story. Students pass if they meet the following criteria:

- They reach their reading rate goal (words correct per minute).
- They make no more than three reading errors.
- They read with good expression.
- They answer the questions correctly.

If they do not pass, teachers assign some remedial activity and test the students again.
Results

Results Screen

When a student passes, the program provides visual and auditory feedback. It also enables the teacher to review the results of the student's progress through the story.

Screen Highlights

1. The difference between the cold timing bar and the pass timing bar shows the student's progress.
2. The teacher can display a results graph of all the stories the student has read in the current level.
3. The teacher can display the student's detailed results for the current story.
4. The student can select a new story to read.
Appendix L

WJ-III Classifications and DIBELS Risk Status
<table>
<thead>
<tr>
<th>Standard Score Range</th>
<th>WJ III Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>131 and above</td>
<td>Very Superior</td>
</tr>
<tr>
<td>121 to 130</td>
<td>Superior</td>
</tr>
<tr>
<td>111 to 120</td>
<td>High Average</td>
</tr>
<tr>
<td>90 to 110</td>
<td>Average</td>
</tr>
<tr>
<td>80 to 89</td>
<td>Low Average</td>
</tr>
<tr>
<td>70 to 79</td>
<td>Low</td>
</tr>
<tr>
<td>69 and below</td>
<td>Very Low</td>
</tr>
</tbody>
</table>

**DIBELS Risk Status – Spring Benchmark Assessment**

<table>
<thead>
<tr>
<th>Risk Status</th>
<th>Score Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>At Risk</td>
<td>0-69</td>
</tr>
<tr>
<td>Some Risk</td>
<td>70-89</td>
</tr>
<tr>
<td>Low Risk</td>
<td>90 and above</td>
</tr>
</tbody>
</table>