

**Student Achievement Data and the Effect of Teacher Content During
Implementation of the Success for All Reading Edge Program
in a Midwestern School District**

Pamela K. Lingelbach
B.A., Park University, 1994
M.S., Park University, 2001

Submitted to the Graduate Department and Faculty
of the School of Education of Baker University
in partial fulfillment of the requirements for the degree

Doctor of Education
in
Educational Leadership

May 3, 2012

Copyright 2012 by Pamela K. Lingelbach

Dissertation Committee

Major Advisor

Abstract

The purpose of this quantitative, non-experimental, causal-comparative study was to analyze the effect of the implementation of the Success for All (SFA) Reading Edge curriculum on Standardized Test for the Assessment of Reading (STAR) Instructional Reading Level (IRL) scores for below and at/above grade level students in three non-Title I middle schools in the Independence Public School District in Independence, Missouri. In addition, the study examined whether receiving reading instruction from teachers certified in different content areas affected the amount of mean IRL score change on regular STAR administration, particularly for students at different proficiency levels.

Data were collected from archived Test Record Reports of 1,470 students enrolled in the three middle schools during the 2007 to 2010 timeframe. Independent one-sample *t* tests and analysis of variance (ANOVA) tests were used for hypothesis testing. Results yielded consistent positive effects for students categorized below-grade level proficiency. Data for at/above grade-level proficiency students was mixed, with two *t* tests revealing significant positive effects, while six of the 8 one-sample *t* tests revealed statistically significant differences with mean IRL scores decreasing. Eight ANOVAs covering specific timeframes were used to determine if there were any significant differences in mean STAR IRL achievement among students receiving reading instruction via the Reading Edge scripted curriculum from teachers in the content areas of communication arts, math, science, and social studies. Main effect results indicated teacher content area did not affect student IRL achievement. Interaction effect results also indicated content area did not affect student IRL achievement for students at varying proficiency levels.

This study supports use of the Reading Edge scripted curriculum package with students who are categorized as below grade-level proficiency and supports the SFA Foundation's suggestion that all certified staff can attain comparable achievement results when providing instruction utilizing the Reading Edge curriculum model. Further research analyzing instructional fidelity and instructors' reading pedagogy background is warranted.

Dedication

This dissertation is dedicated to my Lord, Jesus Christ, who provides for my needs and blesses me abundantly. Among my greatest blessings are my loving family, Ken, Jordan, and Jacob.

Acknowledgments

This dissertation could not have been completed without the support and encouragement of many people. Most importantly, I would like to thank some very special members of my family. My husband, Ken, has been a constant source of support and a keeper of my dreams since before I was even old enough to drive. Thank-you so much for recognizing what was possible and for helping me achieve my goals. My children, Jordan and Jacob, have spent countless hours by my side as I worked to complete this process. I thank them especially for their patience and understanding. My hope is that they, too, will follow their dreams someday. From my earliest childhood years, my grandparents always encouraged me to do my best and to build a reputation of honor and diligent work. Though they are no longer with us, I hope they know how much they are still loved.

I have also been blessed with committed, reflective, and supportive colleagues throughout this journey. I am thankful for the wisdom and encouragement of Jill Whipple, proofreader, mentor, and friend. Additionally, Leah McGee has been by my side teaching me, learning with me, and being a true friend. Thank-you both for the many ways you have enriched my life.

I want to extend sincere appreciation to Dr. Elizabeth Savidge for giving me so many professional opportunities in the past near-decade of our work together. Thank-you for allowing me to take chances, to learn, and to grow. Thank-you too, for being my friend.

I am deeply indebted to Dr. Patricia Bandre' for her countless hours reading and providing feedback on my work. Her ongoing encouragement and helpful guidance have

been invaluable throughout the writing process. It is a rare thing to find someone with whom I can unabashedly share my love of reading, and to discover that in a mentor was a true gift. Sincere thanks are also due to Ms. Peg Waterman, who patiently guided me through the process of learning to speak the language of statisticians, and to Dr. Susan Rogers, who provided wisdom and insight at numerous times along this journey. I appreciate the unique way each of you supported me as a member of my committee, and I thank you for the sacrifices you made on my behalf.

Table of Contents

Abstract	iii
Dedication	v
Acknowledgements.....	vi
Table of Contents	viii
List of Tables	xii
List of Figures	xv
Chapter One: Introduction	1
Background.....	2
Statement of Problem.....	11
Significance.....	13
Purpose of Study.....	15
Delimitations.....	15
Assumptions.....	17
Research Questions	18
Definition of Terms.....	19
Overview of the Methodology	22
Organization of Study.....	23
Chapter Two: Review of Literature	24
Historical Perspective on the Reading “Crisis”	27
National Reading Panel’s background and influence on Federal policy.....	28

No Child Left Behind	31
Success for All Foundation: An Overview	38
The Reading Edge: An overview.....	39
Perspectives on the research.....	43
An overview of the Reading Edge research	48
Program fidelity	48
Reading Edge research still needed	50
Scripted Instruction for Adolescents: Research Needed.....	51
What Makes an Effective Adolescent Literacy/Reading Teacher?	52
Secondary teachers' efficacy and abilities in teaching reading....	56
Teachers view themselves as content specialists, not literacy specialists	59
The importance of teaching in area of certification/licensure	61
Summary	66
Chapter Three: Methods	67
Research Design.....	67
Population and Sample	68
Sampling Procedures	68
Instrumentation	73
Validity and Reliability.....	77
Data Collection Procedures.....	79
Data Analysis and Hypothesis Tests.....	80
Limitations	82

Summary	83
Chapter Four: Results	84
Descriptive Statistics.....	84
Hypothesis Testing.....	85
Summary	116
Chapter Five: Interpretation and Recommendations	118
Study Summary.....	119
Overview of the Problem	119
Purpose Statement and Research Questions	119
Review of the Methodology.....	120
Major Findings.....	121
Findings Related to the Literature.....	122
Conclusions.....	125
Implications for Action.....	125
Recommendations for Future Research	128
Concluding Remarks.....	129
References.....	131
Appendices.....	147
Appendix A. Independence School District At-Risk Committee	
Documents	148
Appendix B. IRB Application	152
Appendix C. IRB Approval	157

Appendix D. Request to Access and Analyze Independence School District
Renaissance Place Archived Data159

Appendix E. Independence School District Approval to Access and Analyze
Renaissance Place Archived Data161

Appendix F. Tukey’s HSD Year 1 Fall to Year 1 Midpoint Timeframe.....163

Appendix G. Tukey’s HSD Year 1 End of Year to Year 2 Fall Timeframe.....165

Appendix H. Tukey’s HSD Year 2 Midpoint to Year 2 Spring Timeframe.....167

Appendix I. Tukey’s HSD Year 2 Spring to Year 2 End Timeframe.....169

List of Tables

Table 1. District September Enrollment Data by Middle School for Three Years.....	5
Table 2. Ethnicity Percentages 2007-2008	6
Table 3. Ethnicity Percentages 2008-2009	7
Table 4. Ethnicity Percentages 2009-2010	8
Table 5. Average Free/Reduced Lunch Rate Data 2007 to 2010	9
Table 6. Data Collection School Years for Cohorts 1-3	17
Table 7. Number of Student IRL Scores Included Per School Per Cohort, 2007 to 2010	69
Table 8. Departmental Teaching Positions Represented Per School 2008 to 2010.....	71
Table 9. Number of Available STAR Scores Per Content Area, 2008 to 2010.....	72
Table 10. Number of Students Scoring Below vs. At/Above Grade Level Per Timeframe.....	73
Table 11. Number of Students Per School, Per Cohort 2007 to 2010	85
Table 12. Cohort Inclusion for Analysis of RQ3 and RQ4.....	86
Table 13. Students in the “Other” Category Each Timeframe.....	87
Table 14. Instructional Reading Level (IRL) Timeframes Used to Obtain Sample Mean Differences for <i>t</i> Tests Used to Address RQ1 and RQ2	88
Table 15. IRL Below Grade Level Proficiency <i>t</i> Testing, ($\alpha = .05$)	89
Table 16. IRL At/Above Grade Level Proficiency <i>t</i> Testing, ($\alpha = .05$).....	94
Table 17. Main Effect for Content Area, Year 1 Fall to Year 1 Midpoint Timeframe....	99
Table 18. Main Effect for Content Area, Year 1 Spring to Year 1 End Timeframe.....	100

Table 19. Sample STAR IRL Means by Content Area.....	102
Table 20. STAR Mean IRL Scores by Content Area and Proficiency Level, Previous Year End to Year 1 Fall	105
Table 21. STAR Mean IRL Scores by Content Area and Proficiency Level, Year 1 Fall to Year 1 Midpoint	106
Table 22. Results of the Content Area by Proficiency Level Interaction Effect	107
Table 23. STAR Mean IRL Scores by Content Area and Proficiency Level, Year 1 Midpoint to Year 1 Spring	108
Table 24. STAR Mean IRL Scores by Content Area and Proficiency Level, Year 1 Spring to Year 1 End.....	109
Table 25. STAR Mean IRL Score by Content Area and Proficiency Level, Year 1 End of Year to Year 2 Fall.....	110
Table 26. Results of the Content Area by Proficiency Level Interaction Effect	111
Table 27. STAR Mean IRL Score by Content Area and Proficiency Level, Year 2 Fall to Year 2 Midpoint.....	112
Table 28. STAR Mean IRL Scores by Content Area and Proficiency Level, Year 2 Midpoint to Year 2 Spring	113
Table 29. Results of the Content Area by Proficiency Level Interaction Effect	114
Table 30. STAR Mean IRL Scores by Content Area and Proficiency Level, Year 2 Spring to Year 2 End.....	115

Table 31. Results of the Content Area by Proficiency Level Interaction Effect,
Year 2 Spring to Year 2 End Timeframe116

List of Figures

Figure 1. U.S. Long-Term Trends in Reading Achievement for Fourth, Eighth, and Twelfth Graders	34
Figure 2. Long-Term Trends in NAEP Average Scores for 13 Year Olds by Race/Ethnicity	36
Figure 3. The Highly Effective Teacher	54
Figure 4. Sample STAR vocabulary in context test item	74
Figure 5. Sample STAR cloze-text, authentic passage item.....	75

Chapter One

Introduction and Rationale

On December 29, 2002, Congress enacted the No Child Left Behind Act of 2001 (NCLB) as the reauthorization of the Elementary and Secondary Education Act of 1965 (NCLB, 2002). The passage of this law, with its requirements regarding the measurement of academic growth, fostered concern about the needs of struggling readers and turned the public eye toward analyzing academic achievement data (Allington, 2006b). This concern was particularly true about the current and ongoing status of the sometimes-elusive skill of reading.

In 2006, United States Secretary of Education, Margaret Spellings, stated, “Reading is the foundation of all learning, a key factor in earning a high school diploma, and a ticket to success in the 21st century” (U.S. Department of Education, 2006, p. 1). It is difficult to predict with certainty what will constitute success in the 21st century; however, it is easy to assume that society will continue to build upon some of the assertions presented in *Reading to Achieve: A Governor’s Guide to Adolescent Literacy* (Bergman & Biancarosa, 2005). This includes the fact that “higher levels of literacy translate into higher earnings... [and]... failure to achieve certain levels of reading, writing, and critical thinking skills in high school narrows employment prospects and limits preparedness for civic participation” (Bergman & Biancarosa, 2005, p. 4). Indeed, many of the fastest-growing professions have increasing literacy requirements for employees. According to Biancarosa and Snow (2004), “...the twenty-five fastest growing professions have far greater literacy demands, while the twenty-five fastest declining professions have lower than average literacy demands” (p. 8).

Acknowledging students' future literacy needs, a school district that finds itself with a reading achievement deficit must institute change to address the problem. Corum, Kepler, Mattson, and Okerstrom (2007) explained, "Leaders are challenged by the daunting task of creating a collaborative school culture which implements a successful and effective literacy approach to increase student achievement" (p. 4). Unfortunately, due to the multi-faceted nature of the issue, which involves instruction, culture, and established practices, no easy solution exists. Sometimes in the face of the overwhelming body of school improvement literature, an educational leader's most effective path seems unclear. However, philosophical differences of opinion, even those stemming from potentially contradictory research, do not remove educators' obligation to implement literacy instruction, particularly for those struggling with the skill. As school leaders make decisions about the instruction they provide, it is important to reflect and evaluate whether the interventions attempted have indeed met the goal of raising student achievement in reading (Clarke, 2002). The district that was the focus of this study faced such a deficit situation, with numerous students reading below grade level. Consequently, district leaders chose to investigate options to provide reading instruction for middle schoolers.

Conceptual Framework and Background

This section provides background about why adolescent reading instruction in a middle school environment has grown to be a nation-wide concern. Second, demographics of the district in the present study, including the number of schools, enrollment numbers in the middle schools, and relevant ethnicity and socio-economic

status percentages are presented. Finally, historical information about reading instruction within the district is summarized.

Research about middle school reading. “Many educators would intuitively agree: Failure is not an option for today’s students—at least not one we would conceivably choose” (Blankstein, 2004, p. 2). However, an analysis of the National Assessment of Educational Progress (NAEP) results regarding reading indicates many students struggle with proficiency in this area. Since 1969, the NAEP has conducted regular, representative assessments to monitor the status of academics in the United States (Institute of Educational Sciences, 2012). A 2011 publication regarding this assessment data explained, “only 29 percent of America’s eighth-grade public school students [met] the NAEP standard of reading proficiency for their grade level” (Alliance for Excellent Education, 2011, p. 1).

One commonly held view was that in the early elementary years, students were learning the mechanics of reading, but as they progressed through school, the focus shifted to reading to learn (Carnine, Silbert, Kame’enui, & Tarver, 2004; Chall, 1995; Jetton & Alexander, 2004). This did not accurately represent the process through which reading development occurred (Alexander & Jetton, 2000). As Tovani (2004) explained, reading tasks become more complex and demanding in nature as students move into secondary grades. The process of learning to read does not end in the elementary grades, but explicit instruction should continue to address ongoing needs as students face new tasks. The cognitive processes involved with reading become increasingly complicated, text often gets lengthier, and students are expected to be able to read independently, comprehend concepts, and apply new knowledge (Tovani, 2000). In addition to

struggling academically, students who continue to lag behind as readers “...find themselves at a serious disadvantage in social settings, as civil participants, and in the working world” (Biancarosa & Snow, 2004, p. 3).

In April 1999, the National Reading Panel (NRP), a group authorized by the United States Congress to investigate available research in the field of reading, issued a press release explaining that the most effective course of reading instruction included a variety of systematic techniques incorporating phonemic awareness, explicit phonics instruction, and guided oral reading (National Reading Panel, 2000). Ede (2006) interpreted this to mean that the NRP supported the use of a scripted curriculum. In the press release, the NRP did not limit the grades for which the instructional approaches might be considered most effective, with the exception of saying that phonemic awareness and explicit phonics instruction were most appropriate in grades K-6 (National Reading Panel, 2000). However, as Allington (2006b) noted, “much of what might prove useful instructionally in first grade is being misapplied to older children and to children having difficulty” (p. 3). Fisher and Ivey (2006) asserted that “few independent research studies have been conducted on popular commercial reading programs, [and] the use of such programs in secondary schools is burgeoning” (p. 181).

The Independence Public School District. The Independence Public School District (ISD) is located in Jackson County, Missouri. The U.S. Census Bureau State and County Quickfacts page estimated the 2010 Independence population as 116,830. The county population was comprised of 65.9% White persons of non-Hispanic origin, Black 22.4%, Hispanic 8.2 %, and other origins 3.5%. The page also listed the median household annual income as \$47,284 (U.S. Census Bureau, 2011). The Missouri

Department of Elementary and Secondary Education’s (DESE) webpage Kids Count data listed the 2008 Jackson County poverty level as 16.4% (2011a). The poverty threshold for a family of four in 2008 was \$22,025 (U.S. Census Bureau, 2011), and to qualify for free/reduced lunch in the 2008-2009 school year, the threshold for a family of four was \$39,220 (Federal Register, 2008).

During the timeframe covered in the present study, students in the ISD were served by 19 elementary schools, 4 middle schools, and 4 high schools. The sample for this study included students in grades six through eight enrolled in the three middle schools that were a part of the district in 2007. The sample did not include a middle school acquired through partial annexation of a neighboring district in 2008 (Hinson, 2010). Table 1 illustrates the enrollment of the three included middle schools during the 2007-2010 school years.

Table 1

District September Enrollment Data by Middle School for Three Years

	School A	School B	School C
2007 – 2008	661	903	825
2008 – 2009	633	896	853
2009 – 2010	607	878	834

Note: Adapted from Department of Elementary and Secondary Education Missouri Comprehensive Data System: Independence 30 (048077), 2011b, Missouri Department of Elementary and Secondary Education. Retrieved from: <http://mcds.dese.mo.gov/guidedinquiry/District%20and%20Building%20Student%20Indicators/Building%20Demographic%20Data.aspx?rp:District=048077&rp:SchoolYear=2011&rp:SchoolYear=2010&rp:SchoolYear=2009&rp:SchoolYear=2008>

DESE district directory information showed that the free/reduced lunch rate of the three schools averaged 47.2% between 2007 and 2010. To receive Title I classification, a school typically had 40% or more of its students' families qualify as low income according to the definition set by the United States Census (U.S. Department of Education, 2011). At the time of this study, none of the middle schools received Title I funding per district decision (E. Savidge, personal communication, April 2011).

Tables 2-4 present each middle school's ethnicity percentages and total enrollment during each of the years represented in the study. The three middle schools had a similar dominant population, with the white, non-Hispanic student group comprising more than 75% of its demographic.

Table 2

Ethnicity Percentages 2007-2008

	School A	School B	School C
White, non-Hispanic	84.5	82.3	78.9
Black	8.9	7.1	10.2
Hispanic	5.9	7.4	7.4
Other	0.4	3.2	3.5

Note: Adapted from Missouri School Directory Independence 30 District Planning Summary Reports: Demographic Data, 2011b, Missouri Department of Elementary and Secondary Education. Retrieved from: <http://mcds.dese.mo.gov/guidedinquiry/District%20and%20Building%20Student%20Indicators/Building%20Demographic%20Data.aspx?rp:District=048077&rp:SchoolYear=2011&rp:SchoolYear=2010&rp:SchoolYear=2009&rp:SchoolYear=2008>

Table 3 shows ethnicity data for each of the three middle schools during the 2008-2009 school year. The three middle schools had a similar dominant population, with the white, non-Hispanic student group comprising more than 75% of its demographic. The overall distribution remained very similar to the previous year.

Table 3

Ethnicity Percentages 2008-2009

	School A	School B	School C
White, non-Hispanic	81.5	77.3	75.6
Black	8.7	9.7	11.6
Hispanic	6.8	8.9	8.3
Other	3.0	4.1	4.5

Note: Adapted from Missouri School Directory Independence 30 District Planning Summary Reports, 2011b, Missouri Department of Elementary and Secondary Education. Retrieved from: <http://mcds.dese.mo.gov/guidedinquiry/District%20and%20Building%20Student%20Indicators/Building%20Demographic%20Data.aspx?rp:District=048077&rp:SchoolYear=2011&rp:SchoolYear=2010&rp:SchoolYear=2009&rp:SchoolYear=2008>

Table 4 shows the ethnicity information for each of the three middle schools during the 2009-2010 school year. The three middle schools had a similar dominant population, with the white, non-Hispanic student group comprising more than 75% of its demographic. Once again, the ethnicity population distribution remained static.

Table 4

Ethnicity Percentages 2009-2010

	School A	School B	School C
White, non-Hispanic	80.6	77.8	75.4
Black	9.9	9.2	14.3
Hispanic	6.8	9.2	7.4
Other	2.7	3.8	2.9

Note: Adapted from Missouri School Directory Independence 30 District Planning Summary Reports: Demographic Data, 2011b, Missouri Department of Elementary and Secondary Education. Retrieved from: <http://mcds.dese.mo.gov/guidedinquiry/District%20and%20Building%20Student%20Indicators/Building%20Demographic%20Data.aspx?rp:District=048077&rp:SchoolYear=2011&rp:SchoolYear=2010&rp:SchoolYear=2009&rp:SchoolYear=2008>

Table 5 presents each school's percentage of students who qualified for free or reduced lunch rates during the study's timeframe. These percentages remained relatively static between 2007 and 2010, and the three schools were similar in their percentage of students qualifying for free/reduced lunch.

Table 5

Average Free/Reduced Lunch Rate Data 2007 to 2010

	School A	School B	School C
2007-2008	50.9	42.9	42.2
2008-2009	54.1	41.0	47.2
2009-2010	51.6	46.5	48.4

Note: Adapted from Department of Elementary and Secondary Education Missouri Comprehensive Data System: Independence 30 (048077), 2011b, Missouri Department of Elementary and Secondary Education. Retrieved from: <http://mcds.dese.mo.gov/guidedinquiry/District%20and%20Building%20Student%20Indicators/Building%20Demographic%20Data.aspx?rp:District=048077&rp:SchoolYear=2011&rp:SchoolYear=2010&rp:SchoolYear=2009&rp:SchoolYear=2008>

History of SFA Program Implementation within the ISD. In October and December 2006, two literacy-focused Teaching and Learning Coaches (TLCs) for the ISD compiled an overview of Standardized Test for the Assessment of Reading (STAR) data to present to a committee analyzing the needs of the district's at-risk learners (See Appendix A). The group analyzing this data was designated as the At-Risk Committee and was comprised of district level administrators, guidance counselors, teachers whose expertise included working with alternative school students or literacy education, and a school board member. When the At-Risk Committee convened, 69% of ISD middle school students were categorized below grade level proficiency according to Renaissance Learning Corporation's STAR Instructional Reading Level (IRL) score (L. Frederick, personal communication, August 2007). The At-Risk Committee, together with middle

school principals, determined that reading deficits were such a widespread concern at the middle level the topic warranted further investigation.

As a result, the district formed a Reading Committee of approximately 30 middle school teachers, central office staff, and district TLCs to evaluate options for reading intervention and instruction programs targeted at adolescent learners. The Reading Committee decided first to evaluate current district practices utilizing a Response to Intervention (RtI) model and to identify areas of weakness within the district. Next, this group researched programs and resources available to purchase and identified characteristics it required to consider a program one of “good quality” (See Appendix A).

In February 2007, the Reading Committee identified the characteristics it considered most important and determined that any approach should mirror the National Reading Panel’s five pillars: phonemic awareness, phonics, vocabulary, fluency, and comprehension (Learning Point, 2004). The group achieved consensus and set the following program criteria as non-negotiable:

- be research-based
- address the needs of a variety of levels of readers, from intense intervention through enrichment
- include ongoing professional development
- include prescriptive, explicit instruction that allows for teacher flexibility
- include phonemic awareness, phonics instruction, vocabulary instruction, fluency building, and components of comprehension strategy instruction (L. Frederick, personal communication, August 2007).

After reviewing multiple program options, in the spring of 2007, the Reading Committee recommended that the district's three middle schools adopt the Success for All (SFA) Foundation's middle school reading instruction curriculum, The Reading Edge. The committee selected the Reading Edge because members believed it met the criteria outlined in its requirements. The ISD adopted and implemented the Reading Edge program in the 2007-2008 school year to provide formal reading instruction to a wide audience of sixth through eighth grade students, ranging from those significantly below to significantly above grade level (See Appendix A).

Statement of Problem

At the time of program implementation for the school district that is the focus of the present study, the majority of the literature on SFA use and student achievement examined the effectiveness of the elementary level SFA reading program or elements of its package (Atkinson, 1998; Blendinger & Wells, 2001; Borman, Slavin, Cheung, Chamberlain, & Madden, 2004; Borman, Slavin, Cheung, Chamberlain, Madden, & Chambers, 2005; Borman, Slavin, Cheung, Chamberlain, Madden & Chambers, 2007; Dianda & Flaherty, 1995; Hurley, Chamberlain & Slavin, 2001; Joyce, 1999; Klingner, Cramer, & Harry, 2006; Ross, Nunnery, Goldfeder, McDonald, Rachor, Hornbeck, & Fleischman, 2004; Slavin, 2002; Slavin & Madden, 2001; Wells, Blendinger & Greene, 2000). However, far less was available regarding the middle school reading curriculum, the Reading Edge. The limited number of studies accessed were largely sponsored or written by individuals affiliated with the SFA Foundation or Johns Hopkins University, which has close ties with the Foundation (Chamberlain, Daniels, Madden, & Slavin, 2007; Daniels, Madden, & Slavin, 2004; Slavin, Daniels, & Madden, 2005). These

studies consistently demonstrated at least limited positive achievement data growth as a result of program usage. However, additional research was essential to evaluate the effectiveness of the SFA Foundation's middle school program, the Reading Edge.

Additionally, in on-site professional development provided to the ISD teachers during initial implementation training, SFA Foundation staff developer Lisa Gaw-Chenausky suggested any certified person was equally qualified to provide instruction using the packaged curriculum materials (personal communication, August 2007). The Reading Edge was structured in such a way that all members of a school's certified staff could possibly have been responsible for teaching a reading class. At the time of the present review of literature, the researcher was unable to locate published evidence to support the assertion that all certified staff should be able to teach reading with equal effectiveness while using a scripted curriculum, warranting further analysis of this program configuration.

The one formal measure of success the ISD used to evaluate the effectiveness of Reading Edge program implementation was the percentage of students considered at/above grade level on the regular administration of the Renaissance Learning STAR assessment and its resulting Instructional Reading Level (IRL) score. For the purposes of the present study, grade level scores were 6.0 IRL for sixth graders, 7.0 IRL for seventh graders, and 8.0 IRL for eighth graders (Alex Ascenvil, Leah McGee, & Lorri Sapp, personal communication, August 2011). Further examination was necessary to determine whether The Reading Edge was a worthwhile intervention, taking into account not only the percentage of students at grade level but also considering measurable score improvement.

Significance

This study provided important information about use of the SFA Foundation's middle school program, the Reading Edge. Multiple studies have been published regarding the elementary SFA reading curriculum, but very few have been conducted about its middle school package. Those published with a middle school emphasis have focused upon a situation in which the program was implemented with a short duration, with compromised fidelity, or with significantly limited grade levels included (Chamberlain, Daniels, Madden, & Slavin, 2007). Although there were few longitudinal studies available about the Reading Edge, Chamberlain, Daniels, Madden and Slavin (2007), researchers closely affiliated with the SFA Foundation, indicated the need for more lengthy investigation. While not longitudinal in nature, the present study covered a three-year time span, analyzing achievement data over two years for three different cohorts of students, a greater span than most studies available at the time of program implementation in the ISD. The findings from this study provided information about the program's effect on academic achievement over a longer time than previously published studies conducted at a district level. Additionally, the majority of publications about the Reading Edge were the work of researchers affiliated with Johns Hopkins University, an institution with close ties to the Foundation, or were the direct work of the SFA Foundation (Pogrow, 2000). In contrast to most available publications, the present study was conducted independently of Johns Hopkins University, the SFA Foundation, and researchers affiliated with either institution, potentially yielding a greater degree of objectivity on the part of the researcher than in previously published research.

Structurally, the Reading Edge program was organized in such a way that a school's entire certified teaching staff was responsible for reading instruction, regardless of their main content area teaching assignment or certification (Slavin, Madden, Chambers & Haxby, 2009). None of the studies located by the researcher explored the relationship between student reading score change and the instructor's content area certification when providing instruction via a scripted curriculum. Consequently, it was important to know whether the students of all certified instructors attained comparable reading achievement change in this teaching situation. Therefore, the findings from this study provided information that facilitates informed decision-making as middle schools consider staffing needs for instruction when utilizing the Reading Edge program.

Another significant aspect of the present study was that the three middle schools represented were non-Title I schools. Title I is a federal program that provides aid to schools serving disadvantaged children and funds reading instruction. On the SFA Foundation's website, the Frequently Asked Questions section addressed this issue by saying:

The majority of SFA schools pay for SFA using Title I and state compensatory education funding *and* many middle class schools successfully use the program *however* most of SFA schools are in high-poverty urban, rural, or inner suburban schools, *where* about 80% of SFA students qualify for free lunch. The greatest obstacle to middle class schools using SFA is financial, since these schools may not have adequate Title I funds or other discretionary funds. However, if schools can find a way to fund the program, SFA works very effectively in middle class schools. (Success for All Foundation, 2011b)

At both the elementary and secondary levels, SFA curriculum implementation is an expensive endeavor for any district choosing to partner with the Foundation.

Historically, the majority of schools utilizing SFA materials have qualified for Title I, which oftentimes helped fund the program. Consequently, the majority of published studies about SFA program effects involved populations of students in Title I schools.

This study provided significant information for the field because it is about Reading Edge program use in a non-Title I environment.

Purpose Statement

The purpose of the present study was to analyze the effect of implementation of the SFA Foundation's Reading Edge program on student reading achievement in three non-Title I middle schools for below and at/above grade level readers as measured by the IRL score on the Renaissance Learning Corporation's STAR test. A second purpose was to examine whether receiving reading instruction from teachers certified in different content areas affected the amount of change on regular academic measures, particularly for students at different proficiency levels.

Delimitations

Delimitations are self-imposed boundaries intended to clarify a study or to narrow its scope (Lunenberg & Irby, 2008). The delimitations utilized by the researcher in the present study were established to provide the maximum amount of complete and consistent data sets available for analysis. Delimitations for this study were as follows:

- The sample of students included in the study was drawn from a single school district in the Midwest and may not be able to be generalized to all middle school situations.

- At the time of Reading Edge implementation in the fall of 2007, only three middle schools were a part of the ISD. A fourth middle school now in the district was excluded from the study because it was not a part of the district until the fall of 2008 when that attendance area was annexed from a neighboring school system. The staff of that middle school had different initial training, and at the time of the present study it was the only district middle school that utilized Title I funding.
- Students had to have nine available STAR test scores within their cohort's designated timeframe for inclusion in this analysis. The study excluded students missing STAR testing data due to transiency or prolonged absence during the test administration date windows.
- All data related to the ISD's use of the Reading Edge program between 2007 and 2010. Within the three years of consistent data available to the researcher, two years of test data were analyzed for each of the cohorts. Cohorts 1 and 2, graduating in 2013 and 2014 respectively, were in grades 7 and 8 during their participation, while Cohort 3, graduating in 2015, participated in grades 6 and 7. For financial reasons, at the end of the third year of program implementation in 2010, all district middle schools changed from a middle school model to a junior high system of organization, significantly affecting a variety of factors related to program implementation and fidelity. Table 6 illustrates the data collection timeframe for each cohort.

Table 6

Data Collection School Years for Cohorts 1-3

	Cohort 1	Cohort 2	Cohort 3
2007-2008	Grade 7		
2008-2009	Grade 8	Grade 7	Grade 6
2009-2010		Grade 8	Grade 7

Note: Each cohort's inclusion covered a two-year timespan, including nine STAR test scores.

Assumptions

Assumptions are the postulates and propositions accepted as operational for the purpose of the study (Lunenberg & Irby, 2008). This study included the following assumptions:

- Teacher effectiveness was equal (communication, classroom management, organization, instructional technique, program fidelity).
- All students exhibited equally diligent effort during reading coursework.
- All students exhibited equally diligent effort on STAR assessments.
- STAR tests were administered in accordance with company suggestions and building and district protocols.
- All instructors received proper initial training, ongoing high-quality professional development, and support to implement the program within its stated guidelines.
- The researcher also assumed all STAR test data obtained from the Renaissance Place database and teacher duty assignments obtained from the PowerSchool student information software were accurate. Additionally, the

researcher assumed all STAR and enrollment data were correctly entered into Microsoft Excel 2010, were properly uploaded into the Statistical Package for the Social Sciences (SPSS) version 19.0, and that the calculations completed through both Microsoft Excel 2010 and SPSS were accurately performed.

Research Questions

Johnson and Christensen (2008) explained, “a research question is a statement of the specific question(s) to which the researcher seeks an answer” (p. 78). Four research questions guided this study:

Research Question 1 (RQ1):

To what extent have mean Instructional Reading Level (IRL) scores on the Renaissance Learning Company’s STAR test improved for students classified as below grade level proficiency during the implementation of the Reading Edge program?

Research Question 2 (RQ2):

To what extent have mean IRL scores on the Renaissance Learning Company’s STAR test improved for students classified at/above grade level proficiency during the implementation of the Reading Edge program?

Research Question 3 (RQ3):

To what extent have mean IRL scores on the Renaissance Learning Company’s STAR test improved for students based upon the content area teaching assignment of their reading teacher?

Research Question 4 (RQ4):

To what extent have mean IRL scores on the Renaissance Learning Company's STAR test improved for students based upon the content area teaching assignment of their reading teacher and their own classification as below or at/above grade level proficiency?

Definition of Terms

The following terms were defined for the purpose of this study:

Assessment. Assessment was defined as determining both formally (through testing) and informally (through questioning) what students have learned and where instruction needs to be adjusted and adapted to achieve mastery (McEwan, 2007).

Below grade level. Below grade level was a level classification for students who did not score at a determined level of proficiency for a grade level. For the purposes of this study, 6.0 IRL was considered grade level for sixth graders, 7.0 IRL for seventh, and 8.0 IRL for eighth (Alex Ascenvil, Leah McGee, & Lorri Sapp, personal communication, August 2011).

Content area teacher. For the purpose of this study, a content area teacher was any teacher currently teaching one of the following subject areas during the majority of teaching hours during the school day: communication arts, math, science, and social studies (Heller & Greenleaf, 2007).

Grade level. Grade level was a score designated as a level of proficiency for reading progress according to a particular grade. For the purposes of this study, grade level scores were 6.0 IRL for sixth graders, 7.0 IRL for seventh graders, and 8.0 IRL for eighth graders. Synonym: proficiency level. (Alex Ascenvil, Leah McGee, & Lorri Sapp, personal communication, August 2011).

Instructional Reading Level (IRL). An Instructional Reading Level score was a “criterion-referenced [measure] generated from the STAR test indicating the highest reading level at which a student is at least 80 percent proficient at recognizing words and understanding material with instructional assistance” (Renaissance Learning, 2010a, p. 41). STAR Reading software determined IRL scores relative to 1995 updated vocabulary lists that are based on the Educational Development Laboratory’s (EDL) A Revised Core Vocabulary (1969). The IRL was estimated based on the student’s pattern of responses to the STAR Reading items. IRL scores were Pre-Primer (PP), Primer (P), grades 1.0 through 12.9, and Post-High School (PHS) (Renaissance Learning, 2010a).

Literacy. Literacy was an individual’s ability to read and write a language (Vacca & Vacca, 2002).

Reading Strategy Instruction. Reading strategy instruction involved explicit transmission of information and modeling focusing upon cognitive mental processes used by readers during the act of reading (McEwan, 2007). Examples of strategies that were commonly accepted within this concept were activating prior knowledge, inferring, clarifying, questioning, predicting, summarizing, and visualizing.

Renaissance Corporation STAR reading test. “The Standardized Test for the Assessment of Reading (STAR) [was] a computer-based multiple-choice reading test designed as a progress-monitoring assessment providing teachers with reading scores for students in grades K–12 in a brief amount of time. Throughout every individual testing session, the database system continually monitor[ed] student responses and use[d] the Adaptive Branching technique to adjust the test difficulty to reflect the skill level of the student” (Renaissance Learning, 2010b). The test format focused upon students giving

correct responses to vocabulary-in-context, text-dependent questions. Students in middle school typically were assessed with 20 vocabulary-in-context items and five passages that asked literal or inferential multiple-choice questions. To answer correctly, students relied on background information, applied vocabulary knowledge, and used active strategies to construct meaning from the text (Renaissance Learning, 2010a).

Scripted curriculum. A scripted curriculum package included instructional materials that were commercially prepared and required the teacher to read from a script while delivering the lesson. They reflected a focus on explicit, direct, systematic skills instruction and attempted to bring uniformity to the quality of instruction that students receive (Ede, 2006).

Success for All Foundation. The Success for All (SFA) Foundation was a nonprofit organization dedicated to the development, evaluation, and dissemination of reform models for preschool, elementary, and middle schools, especially those serving children considered at risk. The SFA Foundation began in 1987 at Johns Hopkins University. As of 2009, the SFA Foundation served about 1,500 schools in 46 states, offering programs in elementary reading, writing, and math, and middle school reading and math at the time of the present study. The Foundation was headquartered in Baltimore, Maryland (Success for All Foundation, 2011a).

The Reading Edge. The Reading Edge was a curriculum product of the SFA Foundation. It was designed for middle school students of diverse backgrounds and achievement levels. The program's goals were to help students become more strategic readers by advancing their word-recognition techniques, improving their reading

comprehension, building their vocabulary, improving their reading fluency, and teaching them to respond to text using the writing process (Robinson, 2002).

Overview of Methodology

In order to address RQ1 and RQ2, three cohorts of students were included, and purposive sampling was used to narrow data collection. A quantitative research design was selected to measure the difference in reading score achievement data between students who measure below or at/above grade level. The researcher obtained data from district-archived STAR test records for participating groups during the spring of 2007 (baseline score) through spring of 2010 and divided student records into three cohorts.

Analysis of RQ1 and RQ2 utilized scores from students in cohorts 1, 2, & 3. A one-sample *t* test was used to evaluate the differences between each of the means and a null value of 0. In addressing RQ3 and RQ4, the researcher categorized teachers within a particular content area based upon their teaching assignment. In addressing RQ3, a two-way analysis of variance (ANOVA) was used to explore differences in achievement among students of teachers of communication arts, math, science, and social studies. Eight quarterly content area mean IRL scores were compared to determine whether there was a significant difference in the academic achievement of students taught by teachers in the four content areas, analyzing the main effect of teaching content area upon student achievement. Analysis for RQ4 utilized the two-factor ANOVA employed for RQ3 to determine whether there was a significant interaction effect between teaching content area and student proficiency level. Score data for RQ1 and RQ2 were analyzed via Microsoft Excel 2010, while RQ3 and RQ4 were compiled and organized in an Excel

spreadsheet then input into the Statistical Package for the Social Sciences (SPSS) version 19.0.

Organization of the Study

The present study is organized in five chapters. Chapter one included the problem statement, background and conceptual framework, study significance and purpose statement, definition of terms, research questions, limitations, delimitations, and assumptions of the study. Chapter two presents a review of literature that includes background about nation-wide trends in reading assessment data and legislation pertaining to reading instruction, an examination of SFA Foundation research base, and an analysis of the relationship between certification and measurable student achievement. Chapter three describes methodology used for this research study. It includes the selection of the participants, data collection, and data analysis procedures as well as an examination of the validity of the instrument used in this study, the Renaissance Learning Corporation's STAR test. Chapter four presents the results of the data analysis for the study's four research questions. Chapter five provides a summary of the study, interpretation of the data and recommendations, conclusions and implications for action and future research.

Chapter Two

Review of Literature

In 1989, The Carnegie Corporation of New York issued *Turning Points: Preparing American Youth for the 21st Century*. This document synthesized 10 years of research in middle level educational practices and was intended to focus reform on middle level education in the United States. As a result of its exploration of best practices, the corporation offered eight principles for improving middle grade education, heavily emphasizing a school configuration of small groups of adolescents taught by expert teachers working closely with one another (Jackson & Davis, 2000). At the heart of the corporation's conceptual framework were the ideal practices of "improving learning, teaching, and assessment for all students" and "creating a culture to support high achievement" (Center for Collaborative Education, n.d, n.p.). Both ideal practices set forth by the corporation clearly had implications for the ways in which schools addressed reading needs and instruction.

In spite of the corporation's decade-old recommendations and the ensuing changes in many schools, a recent National Assessment of Educational Progress (NAEP) assessment showed middle grade literacy scores continued to remain stagnant. The NAEP is a nationwide test periodically given in reading and other subjects to a representative sample of fourth, eighth, and twelfth grade students to assess overall academic achievement (McEwan, 2007). The format of the test is very similar each year, allowing NAEP to document trends and analyze change over time (Institute of Educational Sciences, 2012). The main NAEP test in its current format began in 1990 and has since been administered annually; however, the long-term trend assessment began in 1969 and has only been performed once every four years (Institute of

Educational Sciences, 2012). Scores on either format of the NAEP assessment were based on a representative sample rather than on all 4th, 8th, and 12th grade students in the United States (Institute of Educational Sciences, 2012). A representative sample is designed in such a way that it strongly resembles the overall population and may be used to draw conclusions about it (Johnson & Christensen, 2008).

The National Center for Education Statistics (NCES) long-term trend report showed that in 1994, the average NAEP reading score for 13 year old students was 258; by 2004, it remained an average of 259 (Nation's Report Card, 2011). From 1992 to 2007, the minimum score for proficiency, a performance indicator classifying a student as able to do grade-level work, was 281, far higher than the average score of the representative sample students who were tested. The average 8th grade score for the sample of students tested in both 1994 and 2004 fell within the basic performance range, a level that indicated only partial mastery of the knowledge and skills considered fundamental for a student to complete proficient work within a designated grade (NCES, 2011).

In 2004, the Alliance for Excellent Education published *Reading Next: A Vision for Action and Research in Middle and High School Literacy*, a pivotal document focused specifically upon improving literacy education for adolescents. In it, Biancarosa and Snow (2004) explained that millions of students in the United States lacked the appropriate skills for grade-level reading and listed specific steps for schools to implement in the ongoing quest to teach American youth to read. Among the items they listed as key for effective adolescent literacy development were direct, explicit comprehension instruction, extended time for literacy, and a comprehensive and

coordinated literacy program (Biancarosa & Snow, 2004). As educational reform efforts continue to attempt to develop adequate, appropriate literacy instruction for students in middle grades, educational leaders and lawmakers have a responsibility to a) provide adequate training for pre-service teachers and those already in the field, b) analyze assessment data to identify instructional needs and set meaningful improvement goals, and c) promote collaboration among teachers (Biancarosa & Snow, 2004).

To some extent, the Common Core State Standards (CCSS) Initiative emerged from the aforementioned responsibilities. The CCSS Initiative is an effort to establish shared educational standards among states in the areas of English language arts and mathematics (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). Ideally, this will promote clarity and consistency of expectations from state to state, improving the quality of teacher preparation programs, facilitating collaboration regarding best practices, and enhancing ability to research those practices. Currently, research related to younger elementary students is sometimes applied inappropriately and ineffectively to adolescent instruction (Allington, 2006b). As reforms such as the CCSS Initiative gain momentum, it is important to seek insight into what is effective in the field of middle level education and to evaluate interventions put in place at that level.

This chapter includes research regarding three topics related to the present study. First, background information and data are provided about reading achievement legislation, building a historical framework of understanding about federal expectations regarding reading instruction. Second, the chapter includes a summary of the Success for All (SFA) Foundation's background and philosophy, an explanation of its middle school

reading program, the Reading Edge, and an overview of studies related to the program. Finally, the chapter includes an analysis of secondary instructors as teachers of reading and a review of related literature.

Historical Perspective on the Reading “Crisis”

Concerns about academic achievement in the United States have long been a topic on the political landscape. Lawmakers have grappled with achievement, literacy growth, and keeping the nation’s school children competitive with those from other countries for decades (Federal Education Policy and the States, 2009). One groundbreaking step toward addressing these concerns in the political arena was the National Defense of Education Act (NDEA), a United States Congressional response to the launch of the Sputnik, the first artificial satellite launched into earth’s orbit during the Cold War era by the Soviet Union (Federal Education Policy and the States, 2009). The NDEA was an attempt to bolster the American education system and bring the country to equal footing with Russia in the great Space Race. NDEA was signed into law by President Dwight D. Eisenhower in 1954 and provided funding to support math, science, and modern foreign language instruction in elementary and secondary schools.

Later, in 1965, the Elementary and Secondary Education Act (ESEA) was passed to address social and economic inequalities among schools and students, and it authorized expenditures for “professional development, instructional materials, and resources to support educational programs and parent involvement promotion” (African American Voices in Congress, 2012). The passage of ESEA established Title I, a program designed to provide federal aid to schools serving disadvantaged children and also provide funds to support reading instruction. ESEA has been revised and reauthorized many times since

its first enactment. Allington (2006a) explained that reading instruction funding made available through Title I was supposed to supplement lessons already taking place as a part of regular instruction. However, Title I often replaced regular classroom instruction, which created a situation where many children were pulled for small-group interventions and lessons (Allington, 2006b). Although the intent was to deliver better-designed, more suitable reading instruction to struggling readers, in actual implementation the system was highly fragmented because interventions and lessons were not necessarily supportive of either classroom instruction or an overall school goal (Allington, 2006b). The law's intended design for strengthening and supplementing instruction to provide a solid foundation of skills for all students instead created a system with a myriad of flaws and inconsistencies (Allington, 2006a).

National Reading Panel's background and influence on federal policy. In 1997, Congress authorized the director of the National Institute of Child Health and Human Development (NICHD) to work with the U.S. Secretary of Education to organize a panel to "conduct a comprehensive investigation of research in the field of reading" (Garan, 2002, p. 1). According to a member of the National Reading Panel (NRP), Timothy Shanahan, the group "was charged with determining what research has shown about the effectiveness of instructional approaches, the readiness of these approaches for translation into practice, and the need for further research" (1999, p.1). The 14-member panel conducted a meta-analysis of available reading research. Early in their work, the panel decided to focus upon alphabetic, fluency, and comprehension related topics (National Institute of Child Health and Human Development, 2000). The group gathered information and conducted a series of regional, public hearings that eventually helped

them determine their final areas of focus (National Institute of Child Health and Human Development, 2000). The panel agreed to examine the following issues:

- What impact does phonemic awareness instruction have on learning to read?
- What impact does phonics instruction have on learning to read?
- What impact do repeated readings or other guided oral reading procedures have on learning to read?
- What impact do programs that encourage wide reading have on learning to read?
- What impact does vocabulary instruction have on learning to read?
- What impact does comprehension strategy instruction have on learning to read?
- What impact does teacher education have on learning to read? (Shanahan, 1999, p. 3)

The NRP restricted inclusion in its meta-analysis to those studies that followed evidence-based standards normally associated with research studies in a psychological or medical field (National Institute of Child Health and Human Development, 2000). In choosing to adopt these criteria for including studies, the NRP significantly narrowed the field of available research, because this type of evidence-based research is often organized to examine specific variables in isolation. This study format is often not possible in an educational analysis, because it is difficult to isolate variables from outside circumstances. As a result, many educational studies include a qualitative or mixed-methods approach. Ultimately, 428 studies were included in the NRP's

subcommittee reports, while thousands of studies were excluded because they did not meet the panel's inclusion criteria (Yatvin, 2002).

In addition to the research model utilized by the NRP, other controversies surrounded their work. The published findings for the meta-analysis were quite lengthy, so a summary report was prepared to condense the information. The credibility of the summary report was questioned because it was partially prepared by a public relations firm that was also employed by McGraw-Hill, an educational publishing company with significant financial stake in reading instruction (Strauss, 2002). Casting another shadow over the summary report was the fact panel members acknowledged discrepancies seemed to exist between its data and the ways in which data were communicated in the subcommittee reports (Garan, 2005). Furthermore, some questioned the process by which the quantitative information was extracted for comparative purposes in the subcommittee reports. Camilli and Wolfe (2004) called attention to the fact that in the phonics line of research, the NRP compared studies that “had ‘more active’ phonics interventions (experimental groups) with studies that had ‘less active’ interventions (control groups)...This choice resulted in the experimental groups of some studies resembling the control groups of others” (p. 26). Inconsistencies about the comparisons somewhat weakened the credibility of the Panel's findings and resulting suggestions.

However, NRP members Ehri and Stahl (2001) asserted that critiques published by other members of the panel were highly biased and inaccurate, and noted in particular that Garan's “criticisms [were] highly selected, ill-founded, and overgeneralized” (p. 17). One of the major thrusts of Garan's criticism of the report was its heavy emphasis on

phonics instruction, but Ehri and Stahl (2001) asserted “the NRP findings [were] consistent with the general view that phonics instruction [was] most important in early grades” and went on to explain that the report did not claim to be the “final word on phonics instruction,” particularly noting the report itself posed many questions worthy of study (p. 18). Additionally, they disagreed with the assertion that the measures compared were inconsistent, explaining standardized testing often combines these different elements into an overall reading score. While they conceded Garan’s assertion that some of the writing in the documents was not as clear as one would have hoped, they turned the table to highlight areas of Garan’s criticisms, which they deemed inaccurate (Ehri & Stahl, 2001). The perceived inconsistencies of the published findings of the NRP, coupled with the panel members’ public squabbling about their work on the meta-analysis, undermined their credibility as a conclusive authority about best practices in reading instruction. However, the work of the group was very influential in determining the course of the No Child Left Behind Act (Allington, 2006a).

No Child Left Behind. One of the most significant recent policies governing academic achievement and enacting another revision and reauthorization of ESEA, occurred when President George W. Bush signed the No Child Left Behind (NCLB) Act into law on January 8, 2002. The ESEA revisions adopted as a part of NCLB were heavily based upon the findings and recommendations of the NRP report and have been described as “one of the federal government’s most aggressive attempts to improve elementary and secondary education” (Joftus & Maddox-Dolan, 2003, p.1). This was especially true in the area of funding, because “the federal government severely restrict[ed] the range of reading programs that it [would] fund through NCLB... only

programs that [were] considered to be research-based, using a narrow medical model” (Wilson, Martens, Arya, & Altwerger, 2004, p. 243). Thus, districts that intended to use Title I funding to support reading instruction needed to be cognizant of the information set forth by the NRP report when deciding how to allocate funds.

NCLB was “launched as a bipartisan move for U.S. schooling to address criticisms of lagging literacy standards and a growing achievement gap among students from communities that have historically suffered from socioeconomic disadvantage and cultural marginalization” (Luke & Woods, 2007, p. 202). According to the U.S. Department of Education, Office of Elementary and Secondary Education, it was especially designed to alleviate educational inequities for minority and economically deprived students (2002). However, as noted by Allington (2006), NCLB did not alleviate the fragmentation of reading instructional techniques or interventions created through the original design of Title I. It also did not revolutionize reading instruction, or create a sudden upswing in achievement scores.

Five years after the initial implementation of NCLB, in 2007, students across the United States once again took the annual NAEP exam. In that year’s assessment, the average reading score for 13 year olds was 263, still relatively stagnant for over a decade, (Institute of Educational Sciences, 2012) and only 29 percent of eighth grade students tested scored at the proficient level (Alliance for Excellent Education, 2011). Students who scored at the proficient level were presumed to be competent at completing grade-level appropriate tasks without significant struggle. The low percentage of students who scored proficient was cause for concern because it demonstrated a significant majority of eighth grade students in the nation did not possess the academic skills required to

complete rigorous work considered to be at an eighth grade difficulty level. With only four short years of compulsory education remaining, these students were likely to enter their futures with a crucial deficit.

The 2007 NAEP scores also revealed the startling fact that only 2 percent of all tested eighth graders scored well enough to be deemed at an advanced level (U.S. Department of Education, 2007). While proficiency represented competence at grade-level appropriate tasks, students who were considered advanced were those who demonstrated superior performance on grade-level tasks (NCES, 2011). Students at this level were likely able to read a variety of complex texts, synthesizing the information and completing learning tasks related to the reading (Deschler, Palincsar, Biancarosa, & Nair, 2007). The outcome of the 2007 NAEP assessment indicated schools across the nation were not yet where they needed or wanted to be in terms of academic performance. A continued analysis of NAEP data (see Figure 1) showed scores have remained relatively unchanged at all three assessed grades for 37 years, indicating over three decades during which a majority of the middle school students tested in the U.S. scored below acceptable levels of proficiency.

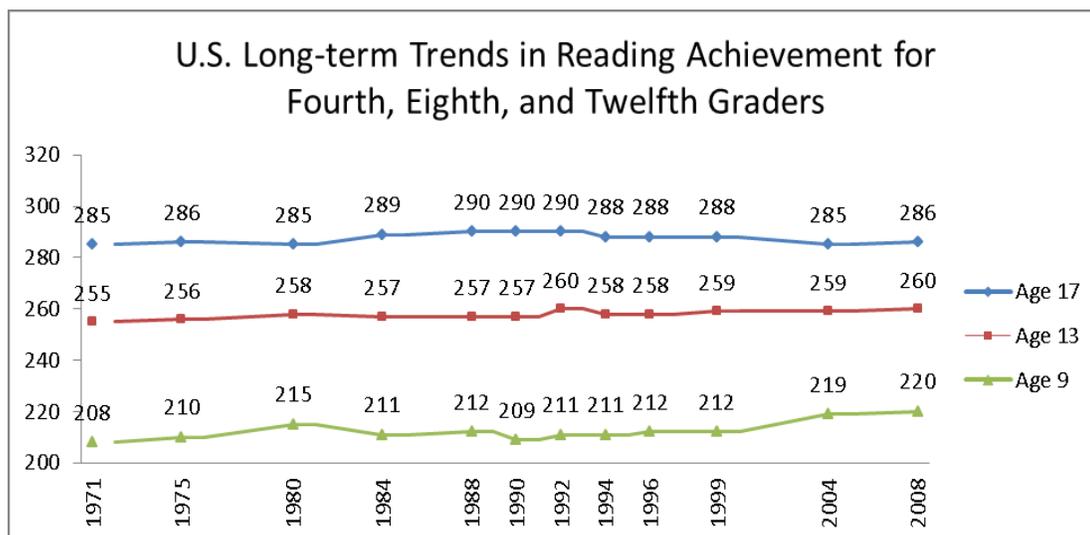


Figure 1. U.S. Long-term Trends in Reading Achievement for Fourth, Eighth, and Twelfth Graders. Adapted from “Fact Sheet” by Alliance for Excellent Education, 2011, *Adolescent Literacy*, Retrieved from <http://www.all4ed.org/files/AdolescentLiteracyFactSheet.pdf>

According to Jennings and Renter (2006), NCLB “intensified the efforts to improve persistently low-performing schools” (p. 4). However, Gallagher (2009) asserted that instead schools now focused “on raising state-mandated reading test scores—the kind of reading our students will rarely, if ever, do in adulthood” (p. 7). A great chasm existed between actually improving reading instruction and student ability and improving their scores on a high-stakes annual assessment. The Alliance for Excellent Education’s 2007 report noted that while great concern existed regarding reading achievement for a significant portion of America’s adolescents, low-income students were far more likely to exhibit this deficit, with the number of students scoring at grade level being as low as one in seven. An additional concern was that evidence pointed to achievement discrepancies among different sub-group populations (Gallagher, 2009).

Trends among different economic groups and other specific populations were closely analyzed on both annual NAEP tests and on annual tests administered at the state level. According to Dennis (2007), ultimately “the goal of NCLB was to help close the achievement gap between disadvantaged students and their peers” (p. 7). Figure 2 shows a score report for 8th graders from 1971-2008, illustrating the different average scores among subgroups over the same timeframe. While overall scores during the timeframe remained stagnant, there was a marginal degree of success in closing the achievement gap.

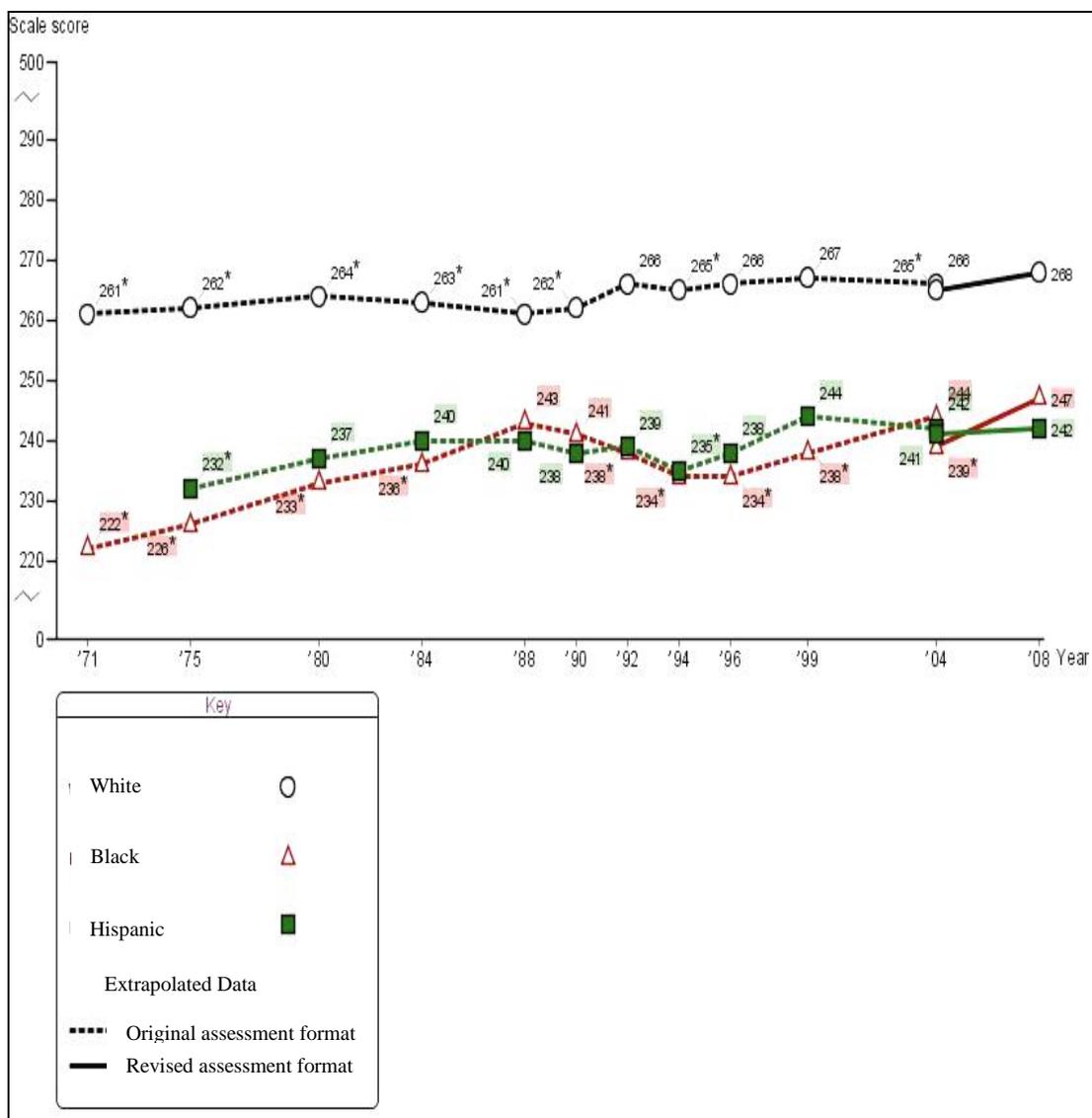


Figure 2. Long-term trend in NAEP average scores for 13 year olds by race/ethnicity

Adapted from “A Nation’s Report Card,” 2012, Retrieved from:

http://nationsreportcard.gov/ltr_2008/ltr0009.asp?subtab_id=Tab_2&tab_id=tab1#chart

NCLB regulations governed not only school-wide achievement levels, but they also monitored the progress of different sub-group populations. Schools faced strict consequences if any of these sub-groups failed to demonstrate adequate growth on assessments (Joftus & Maddox-Dolan, 2003). As a result, the achievement requirements of the NCLB Act coupled with concerns about sub-group progress gaps created a surge of

interest in packaged commercial reading instruction programs as school and district leaders grappled with ways to improve achievement and close gaps (Blanton, Wood, & Taylor, 2007; Duncan-Owens, 2009). “The notion that particular teaching actions can be proven effective is perhaps understandably seductive when millions are being spent on the purchase of reading programs” (Christenbury, Bomer, & Smagorinsky, 2009, p. 10).

Despite the purported intention to alleviate situations of educational inequity, NCLB has been the subject of a great deal of political controversy. From its earliest inception, politicians debated its merit. While some questioned its feasibility, others lauded its philosophical stance (Brown, 2002). The history of party politics aside, NCLB has remained a source of angst among leaders in American education throughout the duration of the present study (Dillon, 2011). In May 2010, the National Education Association (NEA) submitted a lengthy, detailed proposal to Congress, specifying ways in which the law should be amended (NEA, 2011). Prominent among its recommendations was the concept that “ESEA should end the obsession with high-stakes poor quality tests by developing high quality assessment systems that provide multiple ways for students to show what they have learned” (NEA, 2011). A common, related question that has existed since the initial enactment of the law was whether high-stakes assessments were truly an indicator of quality education or were in any way beneficial to students. In a September 2011 speech, U.S. President Barak Obama addressed the issue, explaining that Congress should reform NCLB, acknowledging the higher standards it set forth as an admirable goal, but also stating that in implementation, NCLB had some “serious flaws” (Obama, 2011/2012). The flaws to which he alluded included aspects of funding and testing, two major subjects of contention throughout the law’s life, and two

driving forces in the implementation of packaged, research-based curriculum programs (Obama, 2011/2012). The present study was an analysis of the implementation of one such program available through the Success for All Foundation.

Success for All Foundation: An Overview

The Success for All Foundation (SFA) is a non-profit organization affiliated with Johns Hopkins University and based in Baltimore, Maryland. Development of the SFA instructional program began in 1986 when Baltimore superintendent, Alice Pinderhughes, and school-board president, Robert Embry, requested researchers at Johns Hopkins University design a program to ensure instructional success in schools serving a large population of economically disadvantaged students (Slavin, Madden, Chambers, & Haxby, 2009). The program was first used in September 1987, in a school with a high free or reduced lunch rate and in which over three-fourths of students were African American (Slavin, Madden, Chambers, & Haxby, 2009). In the 1988-1989 school year, the program expanded to five elementary schools within the Baltimore district. Since that time, it has become “one of the leading packaged reading programs on the national market” (Blendinger & Wells, 2001, p. 3) with use in more than 1,200 Title I elementary schools (Borman, Slavin, Cheung, Chamberlain, Madden & Chambers, 2007).

Additionally, 144 middle schools used its middle school curriculum package, the Reading Edge, during the 2006-2007 school year (Chamberlain, Madden, Slavin, 2007).

The SFA program was designed to assist schools with implementing effective instructional practices, organizing school leadership, facilitating parent involvement, regularly assessing students and providing intervention for struggling learners, and implementing a system of ongoing professional development for staff (Slavin, Madden,

Chambers, & Haxby, 2009). According to information available on the Foundation's website, "80% of SFA students qualif[ied] for free lunch" (Success for All Foundation, 2012) and came from economically challenged areas. However, the schools in the present study did not have a population matching this statistic, and were instead a more middle-class environment than an impoverished one. While typical schools that used SFA tended to face severe economic hardship, the Foundation also asserted, "any group of children, regardless of social class, will contain learners with a range of needs, all of which will be accommodated by SFA" (Success for All Foundation, 2012).

The Reading Edge: An overview. The development SFA's middle school program, The Reading Edge, began in 1993 (Lisa Gaw, SFA consultant, personal communication, May 5, 2009). The Reading Edge was described as "a comprehensive literacy program for all middle grade students, from struggling to advanced, with the goal of preparing them to be strategic, independent, and motivated readers and learners" (Slavin, Madden, Chambers, & Haxby, 2009, p. 209). In its *Leadership Guide* training manual, the SFA Foundation described The Reading Edge as a program "designed to help students build a strong phonetic skill base and develop fluency, comprehension strategies, study skills, vocabulary, and a lifelong love of reading" (Success for All Foundation, 2006, p. 6-3). Its texts were selected to be "appropriate to the interests, needs, and developmental characteristics of middle school students" (p. 6-3) and lessons were designed to utilize the inherent social nature of adolescents, heavily emphasizing cooperative learning and interaction.

A 2004 study explained that the SFA middle school reform philosophy took great care to put into place the most important elements of the Carnegie Foundation's 1989

Turning Points report, coupled with an emphasis on strong curriculum and instruction (Daniels, Madden, & Slavin, 2004). While the program was somewhat based on the SFA elementary model, it was revised to meet the developmental needs of adolescents and to accommodate functions within a middle school environment (Daniels, Madden & Slavin, 2004). One of the most notable ways in which the program was re-structured for the middle school environment was the time designated for reading instruction. In the elementary program, a school was expected to set aside a 90-minute block of time for dedicated reading instruction, while at the middle level, this was reduced to 60 minutes, and could successfully be pared down slightly beyond that number.

The Foundation made a few significant procedural and organizational recommendations to schools and districts adopting its Reading Edge program during the 2007-2008 school year. The first recommendation was that all students in the school receive reading instruction during a common time during the day, with the ideal amount of time being a 60-minute block. Because of this schedule structure, most (or all) certified staff members were responsible for teaching a reading class. This facilitated student schedule changes that occurred as a result of demonstrated growth, involved the majority of a staff in literacy related professional development, and promoted building-wide discussion about achievement data.

Another of the Foundation's recommendations was that students be assessed quarterly and re-grouped with students of similar ability to receive ongoing instruction. Within this format, students were grouped into separate class sections according to their demonstrated reading ability levels on both standardized scores and class performance. These groupings fell within three categories: level 1, levels 2 and 3, and levels 4-8+.

Level 1 “focuse[d] on giving beginning readers the tools they need[ed] for literal comprehension. It use[d] a sequence of illustrated stories, presented with phonetically regular text that [became] more difficult as students master[ed] new skills” (Slavin, Madden, Chambers, & Haxby, 2009, p. 6-3). Levels 2 and 3 “use[d] simple fiction, nonfiction, and reader’s theater to focus on basic decoding skills, improving fluency, building vocabulary, and improving comprehension” (Slavin, Madden, Chambers, & Haxby, 2009, p. 6-3). The instructional focus for levels 1, 2, and 3 included basic word attack skills, fluency, and vocabulary development. This focus shifted in levels 4-8+ to encompass more literacy skills and comprehension strategies. All levels had a heavy emphasis on metacognition during the reading process and did so through teacher think-aloud modeling and through discussion prompts for small and large groups.

Daily lesson plans guide[d] teachers to use instructional practices that have been found effective in rigorous research. Among these are cooperative learning (Johnson and Johnson, 1999; Slavin, 1995), metacognitive comprehension strategies (Pressley and Woloshyn, 1995), effective classroom management methods such as rapid pace and active involvement of all students (Evertson, Emmer, and Worsham, 2000). (as cited in Borman, et.al, 2007, p. 704)

Levels 4-8+ materials included a variety of genres, both fiction and non-fiction, to provide lessons with explicit comprehension strategy instruction (Slavin, Madden, Chambers, & Haxby, 2009). The Foundation’s *Overview* booklet outlined the literacy skills covered in each level, from 2-8+. These included cause/effect, compare/contrast, drawing conclusions, identifying main idea, putting events in the proper sequence, identifying fact and opinion, and building increased fluency (Success for All

Foundation, 2004). Curricular materials frequently employed clarifying, summarizing, questioning, and predicting strategies as a part of the lesson and involved explicit instruction about vocabulary, story structure, and literary techniques (Success for All Foundation, 2004).

The Foundation made additional recommendations to districts adopting the Reading Edge curricular program in the 2007-2008 school year. These included a suggestion that each site have a full-time facilitator responsible for maintaining data, overseeing materials, coordinating student placements, and providing ongoing, job-embedded professional development for all program instructors. This professional development was organized in many different formats and for different purposes including but not limited to:

- facilitating group discussions about reading data and instruction
- modeling teaching techniques, planning lessons with instructors
- providing explicit sessions about teaching techniques
- providing feedback about instruction
- providing strategies for formative assessment (L. McGee, personal communication, February 18, 2012).

Follow-up visits were made by the Foundation staff to assist the full-time facilitator and other members of the building leadership team with program implementation issues. These periodic visits consisted of observing classrooms and tutoring sessions; meeting with the facilitator, principal, and leadership team; and providing positive reinforcement and advice regarding a variety of issues (Atkinson, 1998). The job-embedded training provided by a full-time facilitator and the ongoing

feedback from the Johns Hopkins staff provided a mechanism for schools to support ongoing professional development regarding effective reading instructional techniques. Ideally, this helped them become more deliberate in teaching reading in the content areas, allowing students to receive reading craft instruction throughout the day. In today's environment, it has become "critical that secondary content teachers better understand and teach specific literary strategies to help students read and extract meaning from the written material used to teach course content" (NASSP, 2005, p. 1). In 1944, Sterl Artley issued the following injunction to content area teachers, and decades later, his message still rings true:

Every classroom teacher has the direct responsibility for developing those reading skills and abilities essential for adequate comprehension with his particular area of instruction as well as for applying to his content field and making functional those skills and abilities being developed by teachers in other areas of instruction. (as cited in Kamil, 2003, p. 4)

The Reading Edge program did not utilize a reading in the content area approach. Instead, it expected teachers to be competent reading instructors for skill and strategy lessons in general, regardless of content or certification area.

Perspectives on the research. In 1998, Atkinson recommended anyone in the process of evaluating the value and viability of an instructional program or packaged curriculum consider numerous factors. A statistical analysis alone cannot account for all the ways in which students and staff might be impacted by an item of this nature, therefore many aspects of culture and climate should to be taken into consideration (Atkinson, 1998). Additionally, in an educational setting it is almost impossible to

account for all the variables leading to a possible change. A simple structural shift in which reading instruction is implemented where none has previously taken place or where more time is allotted on a daily basis has the potential to cause change in achievement data and student outcomes. “Virtually any reading intervention that reliably increases time engaged in reading should be expected to lead to achievement gains” (Allington, 2006b, p. 46). Likewise, in evaluating the field of available research regarding a program’s statistical success, one must take into account numerous variables, such as replicability of a study, reliability and validity of the study data and methodology, and potential researcher bias.

Research about the elementary curriculum and reform initiatives offered by the SFA Foundation abounds. The typical methodology of these studies was that the program was implemented in a school, that school was matched to a similar control school, and the data were compared. According to Allington (2006b), these studies “generally [found] that the program produce[d] statistically significant achievement” in comparison to the control group (p. 13). A summary of evaluations of studies at the elementary level indicated the SFA reading instructional program did have a significant positive effect on achievement data (Weiler, 1998). When analyzing the published studies of achievement data issued by SFA, one must question the findings in a way similar to which Coles (2001) questioned the publications of the NRP, asking, “Compared to what?” (p. 207). A quantitative study conducted in a school environment has difficulty fully isolating variables that lead to student growth. One must be careful to note there is a great difference between a “comparison of instruction and no instruction, rather than one form of instruction and another” (Coles, 2001, p. 207). In a typical

elementary, one can assume that some direct instruction will take place in a control school, while wise consumers of research will closely read to determine whether that was also the case in studies about middle school implementation. Many middle school schedules do not have time designated specifically for reading instruction.

An overview of Reading Edge research. In 2007-2008, significantly fewer studies were available evaluating the effectiveness of the SFA Foundation's middle school package, the Reading Edge, than were published about its elementary programs. One study that analyzed a yearlong implementation with 405 sixth graders in two rural, high-poverty middle schools used a mixed-methods approach, utilizing both data from quantitative testing and from a significant number of site observations (Chamberlain, Daniels, Madden, & Slavin, 2007). The students were randomly assigned to either a control group, continuing with their existing reading instruction, or to an experimental implementation group. Students in each group completed pre- and post- testing to measure reading gains. The Reading Edge group had higher post-test scores on the vocabulary subscale of the Gates-MacGinitie Reading Test, and the statistically significant difference somewhat established support for use of the model. However, researchers acknowledged that larger, longer studies were needed (Chamberlain, Daniels, Madden, & Slavin, 2007). This was particularly true, as aspects of program fidelity were not possible within the study's structure. Though researchers asserted that teachers in the experimental groups were able to maintain instructional fidelity, classes were scheduled in such a way that students were not regularly evaluated and re-grouped, and there was not a designated program facilitator available to support program instruction and implementation (Chamberlain, Daniels, Madden, & Slavin, 2007). Both of these

components of program usage were heavily emphasized within the elementary literature and were heralded as key items for establishing success (Borman, Slavin, Cheung, Chamberlain, Madden, & Chambers, 2007).

In another study that offered a longer-term analysis of achievement data, researchers analyzed state testing data for students in participating schools versus those in non-participating schools over a three year timeframe, from 2001 to 2004 (Daniels, Madden, & Slavin, 2004). In this study, data related to the SFA middle school program were analyzed by a third-party evaluator, the National Opinion Research Center (NORC). School achievement data from the seven participating and matched comparison schools were pulled from state web sites (Daniels, Madden, & Slavin, 2004). Schools in the study were located in the states of Washington, Missouri, Indiana (two schools), Mississippi, Louisiana, and Arizona. In all seven school pairings, the SFA participating middle schools gained more on the state reading assessments than their matched schools (Daniels, Madden, & Slavin, 2004). Researchers acknowledged the state tests were different measures and had a variety of differing factors to consider. However, when averaging all the achievement data, the student groups in SFA schools gained an average of 24.6 percent on state reading assessments, while their counterparts' average gain was 2.2 percent. State average gains were 4.2 percentage points (Daniels, Madden, & Slavin, 2004).

Sixth, seventh, and eighth grades were represented in the study, but the report did not identify an overall sample size for the student populations included. Demographic data for the participating and matched schools were also omitted, as were the raw scores and proficiency levels for the gains reported. While the design of the study did seem

replicable, these details would strengthen its findings and legitimacy in claiming SFA's Reading Edge program was a worthwhile investment for districts seeking a program to meet the needs of their readers.

Finally, in publications covering both elementary and middle school programs, The Florida Center for Reading Research (2005) noted two strengths of the program to be "instruction is explicit and systematic, and the five components of reading are well-integrated" and "Teacher Editions provide detailed, well-organized and easy to understand lesson plans with detailed directions for teachers" (Robinson, 2002, p. 4). In a Center Report on the Elementary School Comprehensive School Reform Models document (2005), "principals at three SFA schools contacted by the [Comprehensive School Reform Quality] CSRQ Center indicated that the primary advantage of this model [was] that students [were] taught at their instructional level for reading" (p. 227). Though the report focused on the elementary program, the idea of considering text difficulty and complexity was relevant to middle schoolers as well. Allington (2002) acknowledged the importance of matching adolescents with books that were the appropriate level of difficulty within their Zone of Proximal Development (ZPD). The ZPD is defined as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Vygotsky, 1978, p. 86). Utilizing the appropriate level of instruction was one of the program elements noted as a strength in the aforementioned CSRQ report and was a structure in place within both the elementary and middle school programs offered by the SFA Foundation.

Program fidelity. At the elementary and middle school levels, there were common non-instructional program elements to which the Foundation attributed a great deal of importance (Smith, Ross, & Nunnery, 1997). Whether a district implemented a program from SFA as a school reform model or as a specifically targeted reading intervention, the Foundation strongly recommended procedures that acted as a supporting scaffold for the instructional time. These included:

- 1) The program facilitator conducts ongoing, weekly component meetings (professional development) with staff (Borman, Slavin, Cheung, Chamberlain, Madden, & Chambers, 2007).
- 2) The school staff work together with the program facilitator to identify students for targeted, strategy and skill based small group tutoring in addition to the regular classroom instruction (Borman, Slavin, Cheung, Chamberlain, Madden, & Chambers, 2007).
- 3) The school institute an intervention team, known as the Solutions Team, which meets to discuss data, identify needs, and address those needs (Borman, Slavin, Cheung, Chamberlain, Madden, & Chambers, 2007, p. 727). The Solutions Team should include staff, administrators, parents, and community members.
- 4) The school establishes a system for regularly monitoring student attendance and contacting families when concerns arise (Success for All Foundation, 2006).

- 5) The school implements an assessment system that measures students' reading achievement every eight weeks. This is considered a critical element of progress monitoring and the quarterly regrouping process (Success for All Foundation, 2006).

Though all these items were non-instructional, they were designed to establish a structure to provide further support for students and staff. When Foundation trainers conducted site visits to analyze program implementation and discuss data-driven goals, the aforementioned procedures were a part of their fidelity analysis (Success for All Foundation, 2006). In a review of both elementary and middle school SFA programs, researchers noted that in addition to instructional fidelity concerns, the field would benefit from further research about conditions in which the interventions have the most success (CSRQ Center Report on Elementary School Comprehensive School Models, 2005). Often when a scripted program has failed to yield the desired achievement results, it has been attributed to a lack of fidelity in program implementation (Duncan-Owens, 2009, p. 3).

Fidelity of implementation of curriculum reform efforts is often overlooked in evaluations, but it is critically related to how successful a program will be. The degree to which teachers are faithful to a new literacy initiative should therefore be measured to understand why results may be highly variable for one initiative and ensure that comparisons are not muddied by poor or uneven implementation. (Biancarosa & Snow, 2004, p. 26)

Two differing domains of fidelity exist in an analysis of SFA curriculum programming and materials. A primary consideration is how closely teachers adhere to

the instructional lessons and techniques as presented in the printed lesson materials. The Foundation described its curriculum package as research-based lessons and offered training and support from the site-based program coordinator to maintain this fidelity of instructional quality (Success for All Foundation, 2006). Secondly, when considering fidelity, one ought also turn attention to those outside-the-classroom organizational and cultural elements as listed above, as school climate can impact student success (Carnegie Council on Adolescent Development, 1989).

“Bodilly (1996, 1998) and Nunnery (1998) contend that externally developed reforms that are more clearly defined tend to be implemented with greater fidelity and in turn tend to have stronger effects on teaching and learning than reforms that are less clearly defined” (as cited in Borman et al., 2007, p. 707). Many of the instructional techniques in the Reading Edge package were research based, but much research is still needed to determine the effects of non-instructional program modifications in order to accommodate the individual school needs when choosing to utilize this intervention.

Reading Edge research still needed. In addition to the limited availability of studies examining the effectiveness of Reading Edge curriculum materials and related program structures, independently conducted studies were also needed to balance the available field of literature. “The majority of the few published studies [were] authored by the developers and marketers of the materials and programs. In other words, there [were] few independent evaluations of most materials and programs available” (Allington, 2006b, p. 12). Joyce (1999) argued we should not dismiss the studies of the SFA team just because they were conducted by program developers. They were not necessarily biased, but were part of the inquiry into what could be done to address the

national literacy crisis. However, Pogrow (2000, 2001, 2002), a critic of the program, regularly asserted that most studies were conducted by individuals with vested interest in the SFA Foundation or in Johns Hopkins University, and this weakened their credibility in providing an objective analysis.

Researchers affiliated with the Foundation or with Johns Hopkins University also noted areas in which the field would benefit from ongoing research about the Reading Edge program. “Future studies should also continue the interventions over a longer time period, both to determine long-term impacts and to assess program outcomes with teachers who are experienced with the interventions beyond the turbulent early stages of a new program” (Chamberlin & Daniels, 2007, p. 17).

Scripted Instruction for Adolescents: Research Needed

When conducting studies about reading instruction, it can sometimes be difficult to isolate particular variables and attribute achievement differences to them. Chamberlin and Daniels (2007) recognized in their analysis of the middle school SFA package that initial implementation and teacher comfort level may have made a difference in measured achievement. Likewise, in an elementary study, researchers noted, “neither SFA’s scripted nature nor its materials can solve the problems of teachers with weak instructional or management skills. Yet it is hard to separate teacher effects from program effects” (Klingner, Cramer, & Harry, 2006, p. 345).

While elements of the Reading Edge instructional script may have been strongly research-based and may have offered worthwhile lessons on paper, the teacher remained a variable for which researchers needed to account. “There has yet to be a study of the different ways that teachers deliver reading instruction from a script” (Commeyras, 2007,

p. 404). In addition to instructional fidelity, delivery technique is key. A productive classroom environment will always be a critical element to success, regardless of the content addressed. These factors are entirely teacher-dependent and cannot be scripted. In 1999, Moore, Bean, Birdyshaw, and Rycik authored a joint position statement for the Commission on Adolescent Literacy and the International Reading Association. It stated, “adolescents deserve more than a centralized, one-size-fits-all approach to literacy. They deserve teachers who establish productive conditions for learning...” (p. 9). The importance of research to determine how teacher effectiveness influences achievement when delivering lessons from scripted curriculum materials cannot be underestimated.

What Makes an Effective Adolescent Literacy/Reading Teacher?

In 2000, the Board of Directors of the International Reading Association (IRA) adopted a position statement that put forth a list of essential knowledge and practice qualities for an educator to be considered an excellent teacher of reading. These were:

- They understand reading and writing development, and believe all children can learn to read and write.
- They continually assess children’s individual progress and relate reading instruction to children’s previous experiences.
- They know a variety of ways to teach reading, when to use each method, and how to combine the methods into an effective instructional program.
- They offer a variety of materials and texts for children to read.
- They use flexible grouping strategies to tailor instruction to individual students.

- They are good reading “coaches” (that is, they strategically provide help) (p. 1).

These essential qualities remain relevant in almost any reading instructional setting, from entrance into the primary grades through the time students exit at graduation. However, it is important to note that adolescents have social, emotional, and instructional needs unique to their developmental state (Phelps, 2005). Adolescent readers are quite complex, with each having differences in the skills, strengths, background knowledge, and processing ability he or she brings to the classroom. To meet their instructional needs, educators must have the ability to assess and respond to their literacy strengths and areas of deficit (Fisher & Ivey, 2006).

Figure 3 illustrates the instructional practices and behaviors the National Association of Secondary School Principals (NASSP) identified as effective for content area teachers when addressing the combined content and process needs of the learners in their classrooms. Strategic teaching behaviors and instructional practices are both important in establishing an effective adolescent literacy instructional environment. These two differing domains are each addressed in Figure 3.

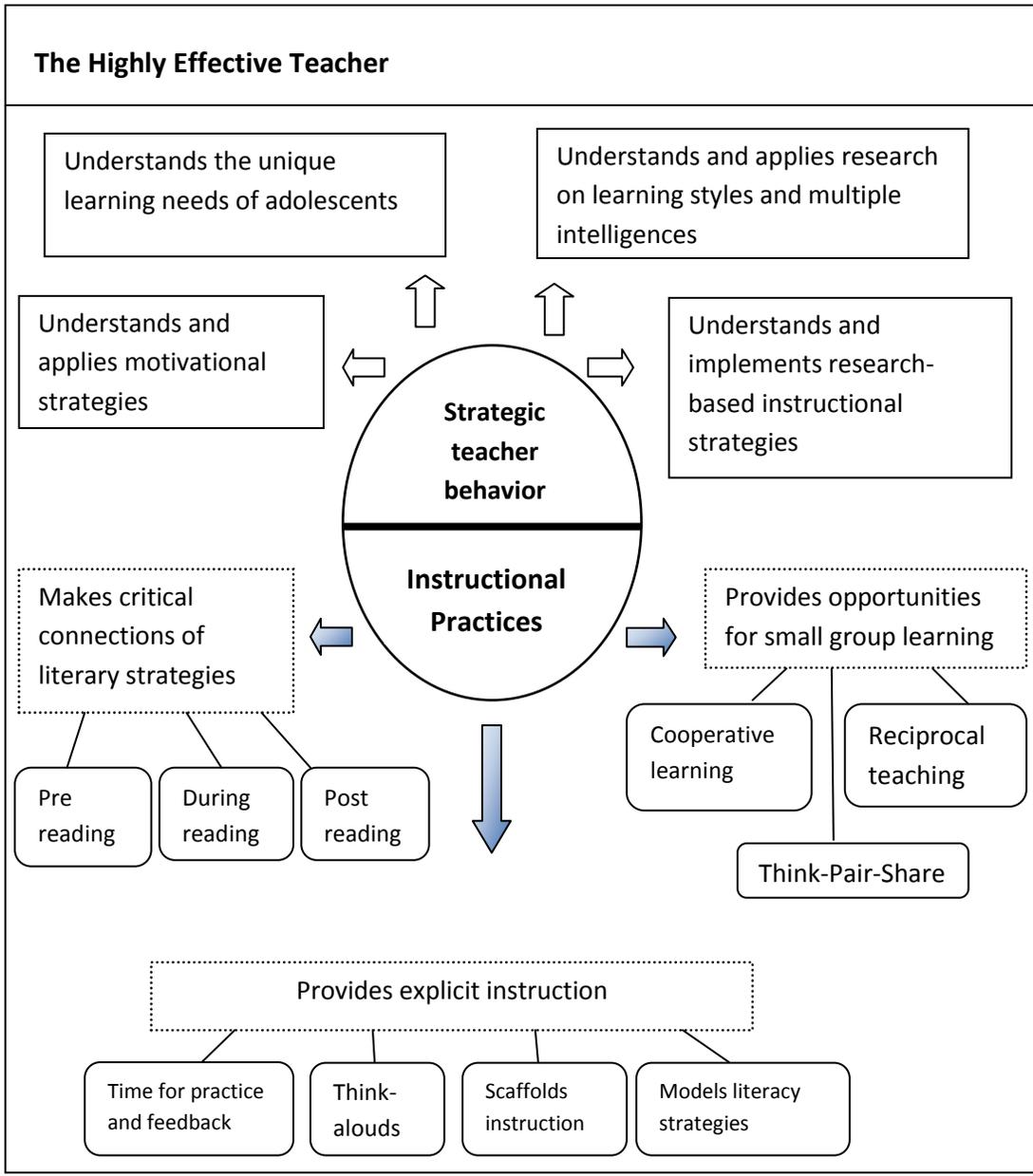


Figure 3. The Highly Effective Teacher. Adapted from *Creating a culture of literacy: A guide to middle and high school principals*, 2005. Reston, VA: National Association of Secondary School Principals.

In addition to the obvious need for expertise in the domain concepts specifically related to a teacher's particular content area, Figure 3 shows identifiable behaviors and practices that can make an individual more effective in teaching the reading skills relevant to his or to her field. These practices span the gulf between merely presenting information and guiding students through the process of learning through reading. Indeed, "one of the most influential things a secondary teacher can do is to use teaching methods that structure reading-based lessons in ways that enable more students to have successful experiences with reading" (Manzo, Manzo, & Thomas, 2005, p. 9).

As Figure 3 depicts, there are two essential elements for effective adolescent literacy instruction. First, the educator must be well-versed in meeting the needs of adolescent learners, strategically implementing research-based methods suitable for engaging and motivating children in this age group (National Council of Teachers of English, 2006; Allington, 2006b). Addressing the varying components ranging from classroom management to approaches that appeal to a variety of learning styles was critical in establishing a climate of learning within the class and promoting literary engagement and growth. Secondly, the educator needs to be adept at a variety of reading instruction techniques to guide students in creating meaning before, during, and after reading sessions. Structures for class configuration may need to vary from individual to small group to whole class scenarios, and these needs are largely dependent upon the particular reading assignments and the abilities of students within the room (Allington, 2006b). An effective teacher monitors all these needs simultaneously, providing explicit instruction appropriate to the task and to the learners.

Defining a highly effective literacy teacher is a complicated, multi-dimensional task. “Literacy is not a subject as such, with a clear disciplinary framework, distinct bodies of knowledge and procedures; but rather a number of disciplines and bodies of knowledge contribute to the content knowledge needed to teach literacy” (Medwell, Wray, Poulson, & Fox, 1998, p. 45). Many variables factor into the equation of each reading lesson, including the task, the goal, the individual abilities of the learners, and the classroom environment. An effective literacy teacher must balance all these components simultaneously.

Although adolescent literacy has grown in prominence, research in the field remains limited, with much of our current understanding based upon studies conducted in the upper elementary grades. Few conclusive studies have examined the variables of effective secondary reading instruction and its translation into growth on indicators of academic achievement (Parris & Block, 2007). Though much is known about effective adolescent instructional techniques in general, the ways in which these methods transfer into a classroom with a literacy or reading in the content area focus are less defined.

Secondary teachers’ efficacy and abilities in teaching reading. “Adolescents entering the adult world in the 21st century will read and write more than at any other time in human history” (Moore, Bean, Birdyshaw, & Rycik, 1999, p. 3). During the past decade, social networking via online text medium and other digital technologies has become a prevalent part of society. The ability to make meaning of the written word and to communicate effectively is no longer relegated to formal exchanges or to being wise consumers of information in the marketplace, as the students of today use writing and the ability to read as a part of an ongoing flow of electronic communication permeating their

lives (Rosen, 2011). They need a strong foundation of literacy skills to do so effectively. Literacy demands in the workplace and as civil participants have also grown at alarming rates. Acquiring a large mental storehouse of factual or procedural information is now far less important than developing the ability to independently research, evaluate information, and communicate findings (Manzo, Manzo, & Thomas, 2005). As social and workplace demands for reading and processing competence increased, a growing body of evidence indicated the need for explicit reading instruction for adolescents. A key element in the effectiveness of such instruction was the quality with which teachers provided it.

Traditionally, education coursework required for secondary certification included few offerings focused upon reading or literacy (National Council of Teachers of English, 2006). Consequently, teachers who were considered highly qualified in their field under the regulations stipulated by NCLB may not have been prepared to address literacy needs even within their own content area (Meyer, 2009; National Council of Teachers of English, 2006). According to a joint position statement from the Commission on Adolescent Literacy and the IRA, “the limited number of reading education courses required for pre-service middle and high school teachers often does not sufficiently prepare them to respond to the escalating needs of adolescent learners” (Moore, Bean, Birdyshaw, & Rycik, 1999, p. 3). In short, secondary instructors have often been ill-prepared or under-prepared to provide necessary reading instruction to support student growth and reading development (Blanton, Wood, & Taylor, 2007; Moore, Bean, Birdyshaw, & Rycik, 1999; NASSP, 2005; NCTE, 2006; Phillips, 2005; Rissman, Miller, & Torgesen, 2009). Indeed, depending upon the requirements of individual

states, a prospective teacher may have only had one reading methods course as a part of his or her teacher preparation coursework (Moats, 1999). This did not provide enough training in the pedagogy of reading and writing instruction skills, nor did it offer enough meaningful opportunities to explore how those practices apply to particular content areas (Heller & Greenleaf, 2007). More fully equipping teachers to address reading and writing concerns in the classroom remains a pivotal factor in developing effective adolescent literacy instruction.

Though the deficit in effective teacher preparation in literacy instruction for secondary educators had been widely acknowledged, little was known about the effect of teacher efficacy on student achievement in reading (Cantrell & Hughes, 2008). The concept of efficacy was first defined by psychologist Alfred Bandura (1977), explaining that it was the belief in “one’s capacity to organize and execute the courses of action required to produce given attainments” (p.3). Teachers are affected by beliefs about their ability to increase student learning, and this in turn influences their behavior as an instructor (Cantrell & Hughes, 2008). The connection between efficacy and expertise in a particular topic seems logical, because common sense indicates that instructors would be most confident and effective in areas in which they have received the most training (Fives, 2003).

In a 2008 study linking teacher efficacy and content-area literacy instruction, Cantrell and Hughes found the teachers who entered the study with the strongest beliefs about their ability to influence students through instruction were more likely to have long-term high levels of skill instruction in their classes (p. 115). As a result, they recommended that:

...because middle and high school teachers often express lower levels of efficacy related to teaching literacy in the content areas, their sense of personal efficacy with literacy teaching should be considered and fostered to promote higher levels of content literacy implementation. (p. 123)

This low level of efficacy regarding teaching reading skills was attributed to the limited training secondary teachers received in this area compared to their more extensive education regarding the content area, and Cantrell and Hughes (2008) suggested that extensive, ongoing professional development should be used to address this concern.

Teachers view themselves as content specialists, not literacy specialists. In a typical secondary school setting, students rotate from one teacher and course to the next in a regular schedule or pattern. This may involve attending class daily or on some other frequency, but regardless of the regularity with which they see their instructors, the norm is for students to be enrolled in the classes of multiple teachers who are responsible for overseeing the varying content areas of communication arts, math, science, social studies. A conventional system of organization within a middle school environment involves arranging students into sets or groups described as a team, and this set of students receives instruction from a common interdisciplinary group of educators (Carnegie Council on Adolescent Development, 1989).

In spite of the fact that such a professional group might have shared responsibility for a designated set of students, and may have worked to establish academic rigor and a safe and nurturing climate, the norm has been for secondary educators to view themselves as content-area specialists, largely to the detriment of literacy skill instruction (Sparks, 2004). Indeed, the responsibility for direct, explicit instruction in developing ongoing

reading and writing skills seemed to belong to no one in particular (Heller & Greenleaf, 2007). Because the academic content of the subject areas has historically been the priority for instruction in communication arts, math, science, and social studies courses, these teachers often expressed beliefs that students should have arrived at their grade level with the necessary reading skills in place. Consequently, many did not consider improving general literacy to be their responsibility (Malette, Henk, Waggoner, & Delaney, 2005).

A variety of factors may have contributed to the content-area specialist mindset of secondary instructors. In addition to the rigors of fast-paced curriculum guides and high-stakes state assessments necessitating judicious allocation of instructional time, teacher efficacy may also have been an issue.

For teachers in middle and high schools, literacy is not, for the most part, an area of expertise. Those who can be described as highly qualified in math, social studies, English [communication arts], or science rarely have any significant training in literacy instruction. (National Council of Teachers of English, 2006, p. 8)

As evidence of this, in a study of certificated secondary content area teachers' knowledge of adolescent literacy, few participants were able to generate a list of basic, widely-accepted comprehension strategies, much less explain their effective use (Meyer, 2009). Meyer (2009) attempted to locate an assessment to measure secondary teachers' knowledge about effective practices in adolescent literacy, but was unable to find an appropriate survey or testing instrument. In the conclusion of the study, Meyer (2009) explained she entered the investigation assuming certificated teachers of adolescents

would have a competent working knowledge of the pedagogy of effective content area literacy instruction. However, study data indicated reading knowledge base of these educators was not as “robust” as anticipated (Meyer, 2009, p. 73). Once again, teachers’ strength seemed to be in their content knowledge and its related training rather than literacy instructional techniques.

The importance of teaching in area of certification/licensure. While commonly accepted as logical that the level of a teacher’s training exerts a positive effect on students’ academic growth, the body of available research did not conclusively link teacher certification and common measures of student achievement (Johnson, 2005). Results from this literature were somewhat contradictory. Some researchers argued that one’s strength of post-secondary education and content knowledge in a field may outweigh the importance of instructional pedagogy, training, and licensure (Goldhaber & Brewer, 2000). Goldhaber and Brewer’s work (2000) comparing the achievement data of students in classes of teachers with traditional licensure versus emergency certification, suggested that a Bachelor of Arts or advanced degree in the content field in which one was teaching may be more important than a teacher’s certification in an area. However, they made no distinction for the reasons people obtained the emergency certification, which left open the possibility some of those with emergency certification had a background of teacher preparation but were new to a content area.

In a meta-analysis of the effect of teacher certification on math, science, and reading achievement, Sparks (2004) bridged the gap between the two somewhat opposing camps of background knowledge importance and pedagogical training. In the areas of math and science, results of the studies examining pedagogy training at elementary,

middle, and high school tended to be somewhat mixed, particularly at middle and high school. Elementary studies indicated a significant effect of certification upon student achievement, but at middle and high school, outcomes were evenly divided between those that did and those that did not indicate an effect of certification in the field. This further validated the assertion of Goldhaber and Brewer (2000) and indicated the importance of background knowledge in a specific content area as opposed to an emphasis on teacher preparation. However, all reading studies included in the meta-analysis strongly pointed toward a positive effect of qualified, certificated teachers providing reading instruction (Sparks, 2004). In this aspect of the meta-analysis, the pedagogy of effective reading instruction was clearly significant.

In contrast to the work of the aforementioned researchers, Denton and Laciara (1982) found a very strong correlative relationship between the amount of teachers' professional training and results on student achievement tests. Hawk, Coble, and Swanson (1985), analyzed the relationship between subject-area certification in mathematics and student achievement and found that in general math and in algebra, scores were significantly affected when students received instruction from appropriately certified staff. Others purported strong evidence existed that training in a traditional teacher preparation program while working toward state licensure yielded superior teachers resulting in students who experienced greater academic achievement and success (Darling-Hammond, Holtzman, Gatlin, & Helig, 2005; Darling-Hammond & Youngs, 2002). In a review of literature, Laczko-Kerr and Berliner (2002) reported that traditionally certified teachers who were providing instruction in an area in which they

were certified had students whose achievement data was higher than that of licensed staff teaching out of their endorsement areas.

Assuming that student achievement scores on assessments are an adequate measure of teacher effectiveness, further study is needed to determine the true impact of certification on teacher quality. Achieving consensus regarding the importance of certification has significant implications for whether districts should hire only those who are fully licensed. Furthermore, when adequately researched to meet the informational needs of the education community, the effects of content knowledge versus certification should be considered when deciding whether instructors should be given teaching assignments outside their fields.

Out-of-field teaching has been defined as a duty assignment in a field or content area outside that of an instructor's endorsement, certification, or licensure (Ingersoll, 2003; Qu & Becker, 2003). An example of this is an individual with a certification only in the area of science teaching a mathematics course. Additionally, out-of-field assignments may be related to grade levels taught. A person who has elementary certification with grade 6 as its upper boundary would be considered out-of-field when teaching a higher grade. The National Center for Educational Statistics (NCES) reported that middle school students were far more likely to receive instruction from an out-of-field teacher than students enrolled in either elementary or high school (Hill, 2011). Many factors may have influenced this likelihood. Middle schools have traditionally been comprised of grade configurations spanning years that may be considered either elementary or secondary in terms of certification. For example, the middle schools in the present study included students in grades 6-8. In the state of Missouri, grade 6 typically

fell within elementary licensing, while grades 7 and 8 were more likely addressed through secondary content area certification or through middle school endorsement (Missouri DESE, 2012). However, certification grade spans have varied from state to state, further clouding the issue of qualification and out-of-area teaching. As a result of this grade organization and the inconsistencies from state to state, the professional qualifications required to teach in a middle school environment were somewhat to blame for out-of-field teaching assignments (Ingersoll, 2003).

One item of importance when considering how student achievement is impacted by out-of-field teaching is the relevance of pedagogical knowledge. Learning to read is a complex achievement, as it is a skill acquired over many years and continually honed through regular practice. The foundational importance of understanding the psychology of reading, reading development, instructional skills and techniques, assessment, and general reading craft cannot be underestimated (Moats, 1999). Teaching any subject is a complex form of mental work involving multiple, rapid decisions that are highly dependent upon the expertise of the instructor. Although the adult in the classroom may be the most expert reader present, it did not necessarily translate into an expertise in the specialized pedagogy needed to provide reading skill and strategy instruction to developing readers (Moats, 1999).

In a meta-analysis that explored research on the effect of teacher certification on reading achievement data for students in grades 3-8, Sparks (2004) presented findings from individual studies and explained overall trends resulting from the comparative data. A primary focus within the meta-analysis was whether teachers were deemed qualified for the instructional task based upon their area of certification. Each study included a

breakdown of data, categorizing teachers as either fully certified or under-certified.

Under-certified teachers were those operating on emergency, temporary, or provisional certification. While these educators were not necessarily out-of-field, their training background was not considered as complete as a fully certified teacher's was. All reading studies included in the meta-analysis found a positive effect of fully certified teachers on student reading achievement. None of the studies included data representing a time in which student achievement was equal to or higher for the students of under-certified teachers than fully certified teachers (Sparks, 2004). The findings of this meta-analysis "provide[d] support for the claim that students of teachers with full certification exhibit higher achievement in reading than students of less-than-fully certified teachers" (Sparks, 2004, p. 83).

Implications from the results of studies analyzing reading achievement data for students of out-of-field teachers and under-certified teachers were significant in an examination of SFA Reading Edge program usage. The majority of instructors utilizing the scripted curriculum package materials in the present study were content area teachers with minimal background in reading pedagogy training. Because they worked in a middle school environment, their certification may have been elementary, with its more thorough reading craft instructional training, or it may have been secondary content focused, with minimal coursework preparing them for this type of instruction. Additionally, those with secondary content-area certification may have been defined as teaching out-of-field when assigned to teach a reading course. Further study is needed to determine whether certification impacts student reading achievement (Darling-

Hammond, Holtzman, Gatlin, & Heilig, 2005), particularly when using a scripted curriculum package (Ede, 2006).

Summary

This review of literature served as an overview of reading legislation and its ongoing impact on current practices. A detailed explanation of the Success for All Foundation's scripted middle school instruction program, the Reading Edge, was also included. Finally, a brief examination of the importance of certification in one's field and the resulting implications for teaching reading concluded this chapter. Next, in chapter three, the researcher discusses the present study's design, population, sample, and sampling procedures. Additionally, information regarding instrumentation, validity, reliability, and data collection is provided. Finally, a description of the study's data analysis, hypotheses, assumptions, and limitations concludes chapter three.

Chapter Three

Methods

The purpose of this study was to analyze the effect of the Success for All (SFA) Reading Edge curriculum implementation on Standardized Test for the Assessment of Reading (STAR) Instructional Reading Level (IRL) scores for below and at/above grade level students in three non-Title I middle schools. In addition, the study examined whether receiving reading instruction from teachers certified in different content areas affected the amount of growth on regular academic measures, particularly for students at different proficiency levels. This chapter presents the methodology used while conducting the research study. It includes a description of the sample population, instrumentation, validity and reliability, data collection techniques, and data analysis procedures.

Research Design

The design of this study was quantitative, non-experimental, and causal-comparative in nature. According to Johnson and Christensen (2008), causal-comparative research exists when “there is one categorical independent variable and one quantitative dependent variable” (p. 360). An independent variable in each of the four research questions was whether students were classified as at/above or below grade level. Students were categorized based on their quarterly STAR IRL scores in comparison with their grade level. To be considered on grade level, sixth graders had to have an IRL score of 6.0, seventh graders 7.0, and eighth graders 8.0. An independent variable in research questions three and four was teacher content area. Differences in instructors to whom students were assigned from quarter to quarter may have influenced score differences for

those classified as either below or at/above grade level, because teachers in some particular content areas may have had more expertise in the pedagogy of reading instruction than others. The quantitative dependent variable within the present study was student IRL change scores on regular STAR assessments. IRL change scores were also used to analyze possible differences in achievement for each quarter-long reading class taught by communication arts, math, science, and social studies teachers during the 2008-2010 school years.

Population and Sample

The population for this study included middle school students, grades six through eight, enrolled in the three schools that were a part of the Independence Public School District from 2007-2010. The sample size for this study consisted of 1,470 students who met the criteria described below.

Sampling Procedures

This study employed purposive sampling in identifying students from the overall population. Johnson and Christensen (2008) described purposive sampling as a situation in which a “researcher specifies the characteristics of the population of interest and locates individuals with those characteristics” (p. 239). Students included in the study were enrolled in three Independence Public School District (ISD) middle schools comprised of grades six through eight during the 2007-2010 school years. In the portion of the study that addressed RQ1 and RQ2, analyzing reading changes on the STAR scores during implementation of the Reading Edge program, students in cohorts 1, 2, and 3 were included if all score data were archived and available to the researcher. To be included in the study, students within a cohort had to have a complete test record of nine

scores available through the Renaissance Learning Company's score database during the appropriate timeframe for the cohort. The graduating class of 2013, referred to as Cohort 1, participated in the Reading Edge program for two years, beginning their seventh grade year. STAR test data for this group included their baseline score from the spring of 2007 through May 2009, which was the group's exit from middle school. The graduating class of 2014 was cohort 2. Their included STAR scores ranged from spring 2008 through May 2010, which marked their exit from middle school. All scores analyzed for this group occurred while these students were in seventh and eighth grades. Cohort 3 was the graduating class of 2015. This group completed two years of participation in the program during their sixth and seventh grade years, and their STAR test data ranged from a baseline in the spring of 2008 prior to program entry through May 2010. Table 7 shows the number of students included from each school in each cohort, divided per school.

Table 7

Number of Student IRL Scores Included Per School Per Cohort, 2007 to 2010

	School A	School B	School C	Cohort total
Cohort 1 (2007-2009)	128	194	198	520
Cohort 2 (2008-2010)	132	209	202	543
Cohort 3 (2008-2010)	130	97	180	407

Note: Information obtained from archived STAR score data.

In the portion of the study that addressed RQ3 and RQ4, analyzing the relationship between student achievement and teacher content area, students' quarterly

mean IRL score changes for the students of communication arts, math, social studies, and science teachers were analyzed. A typical teaching assignment included one common plan time for professional development, one personal plan time, one hour of reading instruction, and four hours in content-area instruction within one or more departments. Teachers were categorized within the previously listed departments if their teaching assignment included three or more periods of instruction in one of those areas during the seven period day. Staff who did not meet the criteria because they did not teach enough hours in the designated content areas, or who taught in completely unrelated content areas, were categorized as “other” and were excluded from the study. For example, an instructor who taught two hours of math and two of social studies was not included in this analysis. The sample for analysis RQ3 and RQ4 was 1,470 students; however, only students enrolled in the classes of communication arts, math, social studies, and science teachers were included in the data analysis. Students enrolled in the classes of instructors deemed as “other” were excluded. To ensure accuracy of enrollment information, only the records of students in cohorts 2 and 3 were included in the data analysis related to RQ3 and RQ4. Table 8 shows the number of teaching positions represented for each department for each school from 2008-2010.

Table 8

Departmental Teaching Positions Represented Per School, 2008 to 2010

	School A	School B	School C
Communication Arts	6	9	9
Math	6	9	9
Science	6	9	9
Social Studies	6	9	9

Note: Information obtained from PowerSchool student information system software.

In the portion of the study that analyzed RQ3 and RQ4, exploring whether a relationship existed between teachers' designated content areas and the changes students exhibited on regular measures of reading, group mean IRL score changes were calculated on a quarterly basis. Table 9 shows the quarterly number of scores used to calculate the mean IRL score changes in content areas. The sample for these calculations included only the students of teachers whose main instructional assignments were communication arts, math, social studies, and science, excluding students of those categorized as "other."

Table 9

Number of Available STAR Scores Per Content Area, 2008 to 2010

Test	Timeframe	Communication Arts	Math	Science	Social Studies
1	Previous Year End to Y1 Fall	257	200	136	240
2	Y1 Fall to Y1 Midpoint	266	214	132	225
3	Y1 Midpoint to Y1 Spring	263	237	131	211
4	Y1 Spring to Y1 End	256	224	133	212
5	Previous Y1 End to Y2 Fall	202	253	184	223
6	Y2 Fall to Y2 Midpoint	202	240	199	213
7	Y2 Midpoint to Y2 Spring	199	245	180	260
8	Y2 Spring to Y2 End	228	246	176	251

Note: Y = Year

Table 10 shows the number of students categorized below grade level and those at/above grade level for the calculation of mean IRL score changes over designated timeframes. The sample for these calculations included the same students as those presented in Table 9.

Table 10

Number of Students Scoring Below or At/Above Grade Level Proficiency Per Timeframe

Test	Timeframe	Below grade level	At/Above grade level
1	Previous year end to Y1 Fall	527	206
2	Y1 Fall to Y1 Midpoint	495	342
3	Y1 Midpoint to Y1 Spring	445	397
4	Y1 Spring to Y1 End	496	329
5	Previous Y1 End to Y2 Fall	489	373
6	Y2 Fall to Y2 Midpoint	486	380
7	Y2 Mid to Y2 Spring	476	378
8	Y2 Spring to Y2 End	452	449

Note: Y = Year

Instrumentation

The STAR is a computer-adaptive reading comprehension test marketed by the Renaissance Learning Company. The company's technical manual explains two of the many purposes of the STAR test use to be: 1) identifying "estimates of reading comprehension using students' instructional reading levels" and 2) tracking reading comprehension growth (Renaissance Learning, 2010a, p. 2).

The STAR test was designed to be administered to large or small groups of students believed to have at least a 100-word sight reading vocabulary as determined by teacher observation and can be completed in a relatively short amount of time (Renaissance Learning, 2010a). The computerized STAR test is given to students in

grades three and up. The test “administers 20 vocabulary-in-context items in the first section of the test and five authentic cloze-format passages with multiple choice literal or inferential questions in the second section of the test” (Renaissance Learning, 2010a, p.

3). Figure 4 represents a sample vocabulary in context item.

Bianca White		1
	I can tie my _____.	
	1.	car
	2.	shoe
	3.	sky

Figure 4. Sample STAR vocabulary in context test item. Adapted from “STAR reading resources: Pretest instructions,” by Renaissance Learning, 2009. Retrieved from <http://renplace.indep.k12.mo.us/RenaissanceServer/SR/Resources/Resources.aspx>

As a student completes the cloze sentence vocabulary-in-context section that comprises the last five questions of the test, he or she must interpret the meaning of the sentence and then select the correct answer from a multiple-choice list of three or four possible answer options. Each of the answer options fits the sentence either semantically or syntactically, so to achieve a correct answer, the student must comprehend the passage as well as know the meaning of the vocabulary word that is the correct answer (Renaissance Learning, 2010a). The student enters an answer selection via mouse or

keyboard (Renaissance Learning, 2010a). Figure 5 represents a sample cloze-format, authentic text passage.

Bianca White		25 / 25
	<p>And, it being a windy day, Little Oliver was completely enshrouded by the skirts of Mr. Bumble's coat as they blew open, and disclosed to great advantage his flapped waistcoat and drab plush knee-breeches. He heard the news of his destination, in perfect silence; and, having had his luggage put into his hand, he pulled his cap over his eyes; and once more attaching himself to Mr. Bumble's coat cuff, was led away by that dignitary. Mr. Bumble drew Oliver along, without notice or remark; for the deacon carried his head very erect, as he always should.</p>	
	1. colleague	2. creditor
	3. dignitary	4. namesake

Figure 5. Sample STAR cloze-text, authentic passage item. Adapted from “STAR reading resources: Pretest instructions,” by Renaissance Learning, 2009. Retrieved from <http://renplace.indep.k12.mo.us/RenaissanceServer/SR/Resources/Resources.aspx>

In both the vocabulary-in-context and authentic passage format questions, “the correct answer option is a word selected from the appropriate grade level of the item set. Incorrect answer choices are words at the same test level or one grade below” (Renaissance Learning, 2010a, p. 13). This difficulty level is determined using the Educational Development Laboratory’s (EDL) Core Vocabulary List which “categorizes

hundreds of vocabulary words according to grade placement, from primer through grade 13 (post high school)” (Renaissance Learning, 2010b, p. 2).

The authentic text portion of the test employs passages “extracted from children’s and young adult literature, from non-fiction books, and from newspapers, magazines, and encyclopedias” (Renaissance Learning, 2010b, p. 4). The Flesch-Kincaid readability estimate is the scale used to rate passages for their difficulty level (Renaissance Learning, 2010b). The readability scores generated from analysis within this system may be converted for academic purposes by using a scale representing the estimated reading difficulty level associated with each grade in school (Dubay, 2004). The authentic passage portion of the STAR test consists of a paragraph analyzed using the Flesch-Kincaid readability scale. In these passages, “the second half of the paragraph contains a sentence with a blank [representing] a missing word” (Renaissance Learning, 2010b, p. 5). The passages in the database range from 27-107 words, but the length of what students encounter during a test session is determined by assessed ability level and by reading speed (Renaissance Learning, 2010b).

Authentic text passages fall within the following categories:

1. Antecedent-consequence: causal relationships are found between sentences.
2. Response: a question-answer or a problem solving format.
3. Comparison: Similarities and differences between sentences are found.
4. Collection: Sentences are grouped together based on some common idea or event. This would include a sequence of events.

5. Description: Sentences provide information by explanation, in specific attributes of the topic, or elaboration on a setting. (Renaissance Learning, 2010a, p. 15).

The STAR reading 4.3 program, which was used for all testing in the present study, had a database with 2,048 test items (Renaissance Learning, 2010a, p. 6). As a student progress through a testing session, the program used adaptive branching to customize the passage difficulty level. Adaptive branching is response-based computer programming designed to “adapt item selection to the examinee’s ability in order to measure as precise[ly] as possible” (Veldkamp, 2010, p. 149). During the STAR test, “if [a] student answer[ed] [an] item correctly, the software bump[ed] up the difficulty of the next item. If the student answer[ed] incorrectly, the software lower[ed] the difficulty of the next item” (Renaissance Learning, 2010c).

Validity and Reliability. The Renaissance Learning Company explained the internal reliability testing of the STAR assessment as it pertains to version 4.3 of the software, saying, “there are three direct methods that can be used to estimate the reliability of the STAR Reading computerized-adaptive test: the split-half method, the test-retest method, and the estimation of generic reliability” (Renaissance Learning, 2010b, p. 11). In a norming study the company conducted in the spring of 2008, using a sample of 69,738 test takers, all internal reliability testing techniques validated use of the software for screening purposes (Renaissance Learning, 2010b). Norm testing results revealed generic reliability estimates to have coefficients ranging from 0.89 in grades 3 and 4 to 0.93 in grades 10, 11, and 12 (Renaissance, 2010b). Because the STAR’s use in

the present study was achievement screening and was not used for diagnostic purposes, it was a valid instrument for the current analysis.

The *Buros Mental Measures Yearbook* included a critique of the STAR test in 2001, analyzing the 2.2 version of the software, which is an earlier, foundational version of the 4.3 software used currently. The review of this earlier version of the software was applicable to the later version because the test structure and format were unchanged. The database of possible questions was larger in the 4.3 version but was inclusive of questions that were in the 2.2 version. The review highlighted the fact that design simplicity and practice items were “easy to use and support[ed] the consistency of administration” (p. 2). Regarding the item bank that was a part of the program at that time, the Buros reviewers found “detailed specifications guided development of all items” (p. 3). Concerns the reviewers noted included that the EDL Core Vocabulary list may “be dated in its presentation of materials relevant to a diverse examinee population” (p. 4). Additionally, they found “the items used to estimate reading ability [were] narrow in focus and remain[ed] heavily influenced by the development of a specific vocabulary” (p. 5). Buros reviewers noted, “The breadth of information obtained is restricted, making diagnostic use of this test limited, which is consistent with the purpose of the assessment as a measure of achievement, not as a diagnostic instrument” (Waterman and Sargent as cited in Nebelsick-Gullet, 2003, p. 6).

In an unpublished dissertation, Benicoff-Nan (2002) explored the use of computer-adaptive branching in order to support or refute the Renaissance Learning company’s assertion that the STAR test could be used as a predictor of achievement on high-stakes testing. In analyzing the correlation between STAR results and California

state testing achievement data for elementary schools, Benicoff-Nan (2002) found moderate to strong results at all grades tested and concluded that overall the STAR assessment was a valid and reliable predictor of student success on state assessments for California students. Additionally, the researcher determined adaptive branching programming was reliable within this testing mechanism and supported the use of STAR as a tool to measure student achievement in reading, particularly for progress monitoring and program evaluation (Benicoff-Nan, 2002).

Data Collection Procedures

An Institutional Review Board (IRB) form was prepared for Baker University for approval prior to data collection. The IRB form requesting this permission is included in Appendix B. Baker University granted the researcher permission to perform this study, as demonstrated in Appendix C. The researcher then contacted Dr. Elizabeth Savidge, Assistant Superintendent of Middle Schools, in February 2012, as is included in Appendix D. District approval to use testing data was granted in the spring of 2012 (see Appendix E).

Students in three schools completed a quarterly computerized Renaissance Learning Company STAR test. Scores were archived in the company's computerized database and stored on the school district's network server during the 2007-2010 school years. The researcher generated Test Record Reports from this database and compiled IRL scores for each testing window for each student in an Excel spreadsheet. A Test Record Report is a document that shows all STAR score data for a student within a designated date range. The date range can be set as a part of the database query.

The individual IRL scores were the dependent variable analyzed to determine the amount of reading achievement growth by groups of students below or at/above grade level. The IRL score is a criterion-referenced measure “indicating the highest reading level at which a student is at least 80 percent proficient at recognizing words and understanding material with instructional assistance” (Renaissance Learning, 2010a, p. 41). After gaining permission, the researcher collected IRL scores archived in the three middle schools’ Renaissance Place STAR testing database. The researcher then examined Test Record Reports, identified students with full sets of data, and compiled all data in an Excel spreadsheet for analysis.

In the portion of the study addressing questions three and four, analyzing the achievement data of students in relation to teachers in different content areas, the researcher determined teacher assignment to a particular content area by analyzing duties as documented in the PowerSchool Administrative scheduling program for each school for the 2008-2010 years. The archived master schedule data in the district’s PowerSchool Student Information Systems software provided teaching assignment/duty information for certified staff during the 2008-2010 school years. Teachers who had three or more instructional periods within a designated content area were categorized within that area when sorting IRL scores for analysis for RQ3 and RQ4. To ensure accuracy of enrollment information, only the records of students in cohorts 2 and 3 were included in the data analysis related to these two questions. Cohort 1 was excluded from this portion of analysis because the researcher did not have access to their enrollment data within the Student Information System software.

Data Analysis and Hypothesis Testing

A one-sample t test was conducted to evaluate the differences between the mean IRL score changes and the null (no change) value of 0 for each hypothesis test. A two-factor analysis of variance was conducted to evaluate the differences among the mean IRL score changes for each three or more group hypothesis test (Johnson & Christensen, 2008).

The research questions and hypotheses are as follows:

RQ1- To what extent have mean Instructional Reading Level (IRL) scores on the Renaissance Learning Company's STAR test improved for students classified as below grade level proficiency during the implementation of the Reading Edge program?

H₀1- Mean IRL scores for students classified as below grade level proficiency improved during the implementation of the Reading Edge program.

Eight t tests were conducted to address RQ1. Each was conducted to evaluate mean IRL change scores for students categorized as below grade level proficiency from eight time periods between 2007 and 2010.

RQ2- To what extent have mean IRL scores on the Renaissance Learning Company's STAR test improved for students classified at/above grade level proficiency during the implementation of the Reading Edge program?

H₀2- Mean IRL scores for students classified above grade level proficiency improved during the implementation of the Reading Edge program.

Eight t tests were conducted to address RQ2. Each was conducted to evaluate mean IRL change scores for students categorized as at/above grade level proficiency from eight time periods between 2007 and 2010.

RQ3- To what extent have mean IRL scores on the Renaissance Learning Company's STAR test improved for students based upon the content area teaching assignment of their reading teacher?

H₀3- There is a relationship between teacher content area and mean IRL change scores.

Eight hypothesis tests were conducted to address Research Question 3 (RQ3).

The two-factor ANOVAs were conducted using mean IRL score changes from eight time periods from 2007 to 2010. The two categorical variables used to group students were their designated proficiency level, being either below or at/above grade level, and the content designation of the teacher from whom they received reading instruction. The two factor ANOVA can be used to test three hypotheses including the main effect for proficiency level, the main effect for teacher content area, and the interaction effect (Proficiency Level X Content Area). The main effect for content area was used to address RQ3.

RQ4- To what extent have mean IRL scores on the Renaissance Learning Company's STAR test improved for students based upon the content area teaching assignment of their reading teacher and their own classification as below or at/above grade level proficiency?

H₀4- There is a difference in the mean IRL change score students of different content area teachers exhibit if students are below grade level proficiency vs. at/above grade level proficiency.

The eight two-factor ANOVAs used to address RQ3 were also used to address RQ4. These two-factor ANOVAs were conducted using mean IRL change scores in

eight time periods from 2007 to 2010 (See Table 13). The two-way interaction effect was used to address RQ4.

Limitations

Limitations are factors that may have an effect on the interpretation of study results and are not under the control of the researcher (Lunenberg & Irby, 2008). In the present study, implementation of the Reading Edge program may not have been the only factor influencing the IRL achievement data analyzed in this study. Other instructional or environmental factors may have impacted the scores.

Summary

Chapter three provided an overview of this quantitative research study. The research questions and hypotheses were outlined in this chapter, and the population, sample, and sampling procedures were described. Additionally, the Renaissance Learning Company's STAR test was explained in detail. In chapter four, the results of the hypothesis tests are outlined to determine the extent to which the implementation of the SFA Foundation's Reading Edge program impacted student achievement for below and at/above grade level students and the students of teachers whose main teaching duty included the content areas of communication arts, math, social studies, and science.

Chapter Four

Results

The purpose of this study was to analyze the effect of implementation of the Reading Edge program on student reading achievement as measured by Instructional Reading Level (IRL) scores on the Renaissance Learning Company's Standardized Test for the Assessment of Reading (STAR) test. This study was conducted to determine whether there was a statistically significant difference in reading score change on a quarterly basis for students who were categorized at proficiency levels of either below grade level or at/above grade level. A second purpose of the study was to explore the relationship between measured reading change of students below or at/above grade level and instructors' content teaching areas. This analysis was conducted to determine whether students of instructors in a particular content area demonstrated greater change than other students did. The current chapter provides results from the quantitative data analysis used to address the four stated research questions. The findings are presented beginning with an explanation of the descriptive statistics followed by hypothesis testing results.

Descriptive Statistics

The population for the present study included students in sixth, seventh, and eighth grades enrolled in three middle schools in the Independence School District (ISD) between 2007 and 2010. The sample was $N = 1,470$ students who had nine STAR test records available during the appropriate years of representation for their respective cohorts. School A was represented by 390 students; School B had 502 students; and

School C had 578 students included. The number of students per school, per cohort is illustrated in Table 11.

Table 11

Number of IRL Scores Included Per School Per Cohort, 2007 to 2010

	School A	School B	School C	Cohort total
Cohort 1 (2007-2009)	128	194	198	520
Cohort 2 (2008-2010)	132	209	202	543
Cohort 3 (2008-2010)	130	97	180	407

Note: Information obtained from archived STAR score data.

Hypothesis Testing

Microsoft Excel 2010 was used to analyze data for Research Question 1 (RQ1) and Research Question 2 (RQ2), and the IBM SPSS Statistics 19.0 Faculty Pack for Windows program was used to analyze the data for Research Question 3 (RQ3) and Research Question 4 (RQ4). Data for all students in the sample were used for statistical analysis for RQ1 and RQ2. However, not all students in the sample were included in hypothesis testing for RQ3 and RQ4.

In the district's organization of the reading program, some students were taught by faculty members who were not categorized as a core content area teacher, due to the fact that they taught special education, exploratory courses, or other courses as their main teaching assignment. For data analysis purposes, when scores were compiled regarding quarterly instructor assignment, students of staff other than communication arts, math,

science, or social studies teachers were categorized as “other.” There was not significant commonality among their teaching duties to warrant creating another category for analysis, so students taught by those with this instructor designation were excluded on a quarter-by-quarter basis in hypothesis testing for RQ3 and RQ4. Additionally, to ensure accuracy of teacher categorization via records available in the PowerSchool student information system software, only students in cohorts 2 and 3 were used as the sample for content area analysis, excluding one student whose enrollment records were not accessible. Cohort 1 was excluded from this portion of analysis because the researcher did not have access to their enrollment data within the Student Information System software. Table 12 represents the sample of students available for testing RQ3 and RQ4 prior to the quarterly exclusion of students enrolled in classes of instructors designated as “other.”

Table 12

Cohort Inclusion for Analysis of RQ3 and RQ4

	School A	School B	School C	Cohort total
Cohort 2 (2008-2010)	132	209	202	543
Cohort 3 (2008-2010)	130	97	179*	406

Note: Information obtained from archived STAR score data.

* The researcher was unable to obtain complete enrollment information for one student in cohort 3, school c, excluding this individual’s records from analysis in RQ3 and RQ4.

Table 13 illustrates the number of students categorized “other” each of the eight tests. These students were excluded from all statistical analysis for RQ3 and RQ4.

Table 13

Students in the “Other” Category Each Timeframe

Test	Timeframe	Frequency
1	Previous Year End to Y1 Fall	636
2	Y1 Fall to Y1 Midpoint	632
3	Y1 Midpoint to Y1 Spring	627
4	Y1 Spring to Y1 End	644
5	Previous Year 1 End to Y2 Fall	607
6	Y2 Fall to Y2 Midpoint	615
7	Y2 Midpoint to Y2 Spring	627
8	Y2 Spring to Y2 End	568

Note: Y= Year

Eight one-sample t tests were used to analyze mean IRL change scores for students at the proficiency levels of below or at/above grade level for both hypotheses one (H1) and hypothesis two (H2). Each quarter of enrollment during a student’s participation in the program and subsequent inclusion in the present study, the individual took a STAR test, was designated as below or at/above grade level, and possibly changed to a different teacher’s class. As a result of this program configuration and potential student proficiency status changes (moving from below to at/above grade level or vice versa), each quarter’s data were analyzed as a discrete timeframe set, resulting in eight t tests per RQ1 and RQ2. The hypothesis tests were conducted using the change in mean

IRL scores, beginning with the Year 1 Baseline score to the Year 1 Fall timeframe. Table 14 details timeframes used to determine the mean changes for each of the eight *t* tests.

Table 14

Instructional Reading Level (IRL) Timeframes Used to Obtain Sample Mean Differences for t Tests Addressing RQ1 & RQ2

Test	Timeframe	Pre-test administration	Post-test administration
1	Previous Year End to Y1 Fall	Y1 May prior to program entry	Y1 Fall End of first quarter
2	Y1 Fall to Y1 Midpoint	Y1 Fall End of first quarter	Y1 Midpoint End of second quarter
3	Y1 Midpoint to Y1 Spring	Y1 Midpoint End of second quarter	Y1 Spring End of third quarter
4	Y1 Spring to Y1 End	Y1 Spring End of third quarter	Y1 End End of fourth quarter
5	Previous Y1 End to Y2 Fall	Y2 Previous year end of fourth quarter	Y2 Fall End of first quarter
6	Y2 Fall to Y2 Midpoint	Y2 Fall End of first quarter	Y2 Midpoint End of second quarter
7	Y2 Midpoint to Y2 Spring	Y2 Midpoint End of second quarter	Y2 Spring End of third quarter
8	Y2 Spring to Y2 End	Y2 Spring End of third quarter	Y2 End End of fourth quarter

Note: Y = Year

RQ1: To what extent have mean Instructional Reading Level (IRL) scores on the Renaissance Learning Company's STAR test improved for students classified as below grade level proficiency during the implementation of the Reading Edge program?

H1: Mean Instructional Reading Level (IRL) scores for students classified as below grade level proficiency improved during the implementation of the Reading Edge program.

Eight one-sample t tests of mean IRL scores were conducted to address RQ1. For each of the eight t tests, the sample means were tested against the null value of 0 and at the significance level of $\alpha = .05$. Each of the eight t tests compared variables of mean IRL score change for students below grade level proficiency. The results of the t tests for below grade level proficiency students are included in Table 15.

Table 15

IRL Below Grade Level Proficiency t Testing, ($\alpha = .05$)

Test	Timeframe	N	Mean Change	SD	t -statistic	p -value
1	Previous Year End to Y1 Fall	1007	0.354	1.205	9.324	0.000
2	Y1 Fall to Y1 Midpoint	965	0.522	1.251	12.971	0.000
3	Y1 Midpoint to Y1 Spring	931	0.223	1.509	4.516	0.000
4	Y1 Spring to Y1 End	839	0.606	1.285	13.649	0.000
5	Previous Y1 End to Y2 Fall	925	0.632	1.487	12.930	0.000
6	Y2 Fall to Y2 Midpoint	852	0.585	1.349	12.662	0.000
7	Y2 Midpoint to Y2 Spring	828	0.514	1.441	10.259	0.000
8	Y2 Spring to Y2 End	778	0.668	1.522	12.241	0.000

Note: Y = year

The first one-sample t test was conducted to compare the mean IRL score change of students categorized as below grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained the Previous Year End from that of Year 1 Fall. The results of the test ($t = 9.324, \alpha = .000, df = 1,006$) revealed a statistically significant difference. On average the below proficiency level students increased their scores by $M = .354$. This mean IRL score change may indicate support for H1, demonstrating IRL achievement for students below grade level proficiency improved during implementation of the Reading Edge program.

A second one-sample t test was conducted to compare the change in the mean IRL score of students categorized as below grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 1 Fall from that of Year 1 Midpoint. The results of the test ($t = 12.971, \alpha = .000, df = 964$) revealed a statistically significant difference. On average the below proficiency level students increased their scores by $M = .522$. This mean IRL score change may indicate support for H1, demonstrating IRL achievement for students below grade level proficiency improved during implementation of the Reading Edge program.

A third one-sample t test was conducted to compare the change in the mean IRL score of students categorized as below grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 1 Midpoint from that of Year 1 Spring. The results of the test ($t = 4.516, \alpha = .000, df = 930$) revealed a statistically significant difference. On average the below proficiency level students increased their scores by $M = 0.223$. This mean IRL score change may indicate

support for H1, demonstrating IRL achievement for students below grade level proficiency improved during implementation of the Reading Edge program.

A fourth one-sample t test was conducted to compare the change in the mean IRL score of students categorized as below grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 1 Spring from that of Year 1 End of year. The results of the test ($t = 13.649, \alpha = .000, df = 838$) revealed a statistically significant difference. On average the below proficiency level students increased their scores by $M = 0.606$. This mean IRL score change may indicate support for H1, demonstrating IRL achievement for students below grade level proficiency improved during implementation of the Reading Edge program.

A fifth one-sample t test was conducted to compare the change in the mean IRL score of students categorized as below grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained the previous Y1 End from that of Year 2 Fall. The results of the test ($t = 12.930, \alpha = .000, df = 925$) revealed a statistically significant difference. On average the below proficiency level students increased their scores by $M = 0.632$. This mean IRL score change may indicate support for H1, demonstrating IRL achievement for students below grade level proficiency improved during implementation of the Reading Edge program.

A sixth one-sample t test was conducted to compare the change in the mean IRL score of students categorized as below grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 2 Fall from that of Year 2 Midpoint. The results of the test ($t = 12.662, \alpha = .000, df = 851$) revealed a statistically significant difference. On average the below proficiency level students

increased their scores by $M = 0.585$. This mean IRL score change may indicate support for H1, demonstrating IRL achievement for students below grade level proficiency improved during implementation of the Reading Edge program.

A seventh one-sample t test was conducted to compare the change in the mean IRL scores of students categorized as below grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 2 Midpoint from that of Year 2 Spring. The results of the test ($t = 10.259$, $\alpha = .000$, $df = 827$) revealed a statistically significant difference. On average the below proficiency level students increased their scores by $M = 0.514$. This mean IRL score change may indicate support for H1, demonstrating IRL achievement for students below grade level proficiency improved during implementation of the Reading Edge program.

An eighth one-sample t test was conducted to compare the change in the mean IRL scores of students categorized as below grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 2 Spring from that of Year 2 End of year. The results of the test ($t = 12.241$, $\alpha = .000$, $df = 777$) revealed a statistically significant difference. On average the below proficiency level students increased their scores by $M = 0.668$. This mean IRL score change may indicate support for H1, demonstrating IRL achievement for students below grade level proficiency improved during implementation of the Reading Edge program.

In summary, each of the eight one-sample t tests for students scoring below grade level proficiency revealed a significant difference with mean score change. The greatest difference appeared to occur at the conclusion of each of the two years, Year 1 Spring to

Year 1 End ($M = .632$) and Year 2 Spring to Year 2 End ($M = .668$). These increases indicate academic growth as measured by the STAR IRL.

As with the hypothesis tests that were conducted to address RQ1, the hypothesis tests for RQ2 tested the sample means against the null value of 0 and at the significance level of $\alpha = .05$. See Table 14 for a list of testing sessions used to determine the sample means for each t test for RQ2.

RQ2: To what extent have mean IRL scores on the Renaissance Learning Company's STAR test improved for students classified at/above grade level proficiency during the implementation of the Reading Edge program?

H2: Mean IRL scores for students classified above grade level proficiency improved during the implementation of the Reading Edge program.

Eight one-sample t tests of mean IRL score change were conducted to address RQ2. For each of the eight t tests, the sample means were tested against the null value of 0 and at the significance level of $\alpha = .05$. Each of the eight t tests compared variables of mean IRL score change for students at or above grade level proficiency. The results of the hypothesis t tests for at/above grade level proficiency students are included in Table 16.

Table 16

IRL At/Above Grade Level Proficiency t Testing, ($\alpha = .05$)

Test	Timeframe	N	Mean Change	SD	t-statistic	p-value
1	Previous Year End to Y1 Fall	463	-0.391	1.833	-4.589	0.000
2	Y1 Fall to Y1 Midpoint	505	-0.237	1.740	-3.067	0.002
3	Y1 Midpoint to Y1 Spring	539	0.243	1.420	3.979	0.000
4	Y1 Spring to Y1 End	631	0.202	1.702	-2.975	0.003
5	Previous Y1 End to Y2 Fall	543	-0.408	1.739	-5.463	0.000
6	Y2 Fall to Y2 Midpoint	616	-0.233	1.574	-3.672	0.000
7	Y2 Midpoint to Y2 Spring	640	-0.307	1.488	-5.221	0.000
8	Y2 Spring to Y2 End	690	-0.196	1.515	3.400	0.001

Note: Y= Year

The change was calculated by subtracting the mean obtained the previous Year End from that of Year 1 Fall. The results of the first test ($t = -4.589$, $\alpha = .000$, $df = 462$) revealed a statistically significant difference. On average the at/above grade level proficiency level students' mean IRL scores decreased by $M = -0.391$. This negative change in mean IRL scores may provide evidence to reject H2.

A second one-sample t test was conducted to compare the change in the mean IRL scores of students categorized as at/above grade level proficiency as compared to the null

value of 0. The change was calculated by subtracting the mean obtained Year 1 Fall from that of Year 1 Midpoint. The results of the test ($t = -3.067$, $\alpha = .002$, $df = 504$) revealed a statistically significant difference. On average the at/above grade level proficiency students' mean IRL scores decreased by $M = -0.237$. This negative change in mean IRL scores may provide evidence to reject H2.

A third one-sample t test was conducted to compare the change in the mean IRL scores of students categorized as at/above grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 1 Midpoint from that of Year 1 Spring. The results of the test ($t = 3.979$, $\alpha = .000$, $df = 538$) revealed a statistically significant difference. On average the at/above grade level proficiency students increased their mean IRL scores by $M = 0.243$. This mean IRL score change may indicate support for H2, demonstrating mean IRL improvement for students at or above grade level proficiency during implementation of the Reading Edge program.

A fourth one-sample t test was conducted to compare the change in the mean IRL scores of students categorized as at/above grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 1 Spring from that of Year 1 End of year. The results of the test ($t = -2.975$, $\alpha = .003$, $df = 630$) revealed a statistically significant difference. On average the at/above grade level proficiency students' mean IRL scores decreased by $M = -0.202$. This negative change in mean IRL scores may provide evidence to reject H2.

A fifth one-sample t test was conducted to compare the change in the mean IRL scores of students categorized as at/above grade level proficiency as compared to the null

value of 0. The change was calculated by subtracting the mean obtained the previous Y1 End from that of Year 2 Fall. The results of the test ($t = -5.463$, $\alpha = .000$, $df = 543$) revealed a statistically significant difference. On average the at/above grade level proficiency students' mean IRL scores decreased by $M = -0.408$. This negative change in mean IRL score may provide evidence to reject H2.

A sixth one-sample t test was conducted to compare the change in the mean IRL scores of students categorized as at/above grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 2 Fall from that of Year 2 Midpoint. The results of the test ($t = -3.672$, $\alpha = .000$, $df = 615$) revealed a statistically significant difference. On average the at/above grade level proficiency level students' mean IRL scores decreased by $M = -0.233$. This negative change in mean IRL scores may provide evidence to reject H2.

A seventh one-sample t test was conducted to compare the change in the mean IRL scores of students categorized as at/above grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 2 Midpoint from that of Year 2 Spring. The results of the test ($t = -5.221$, $\alpha = .000$, $df = 639$) revealed a statistically significant difference. On average the at/above grade level proficiency students' mean IRL scores decreased by $M = -0.307$. This negative change in mean IRL score may provide evidence to reject H2.

An eighth one-sample t test was conducted to compare the change in the mean IRL scores of students categorized as at/above grade level proficiency as compared to the null value of 0. The change was calculated by subtracting the mean obtained Year 2 Spring from that of Year 2 End of year. The results of the test ($t = 3.40$, $\alpha = .000$, $df =$

689) revealed a statistically significant difference. On average the at/above grade level proficiency students' mean IRL scores decreased by $M = -0.196$. This negative change in mean IRL scores may provide evidence to reject H2.

In summary, each of the eight one-sample t tests conducted for students categorized at/above grade level proficiency revealed a significant difference between mean IRL scores. Six of the eight one-sample t tests revealed statistically significant differences with mean IRL scores decreasing. The greatest decrease appeared to occur at the beginning of the second year of implementation, from Previous Year 1 End to Year 2 Fall ($M = -0.408$). The third one-sample t test revealed a significant difference that indicated a mean IRL score increase for this group.

RQ3: To what extent have mean IRL scores on the Renaissance Learning Company's STAR test improved for students based upon the content area teaching assignment of their reading teacher?

H3: There is a relationship between content area and mean IRL change scores.

Eight two-factor ANOVAs were conducted to address RQ3. The two categorical variables used to group the students' scores were proficiency level (below and at/above grade levels) and teachers' content area (communication arts, math, science, and social studies). A two-factor ANOVA can be used to test three hypotheses including a main effect for proficiency level (below or at/above grade level), a main effect for content (communication arts, math, science, and social studies), and a two-way interaction effect between content and proficiency level. The main effect for content was used to address RQ3, using changes in mean STAR IRL scores as the dependent variable. Each of the

eight ANOVAs was calculated using mean STAR IRL scores from cohorts 2 and 3, and each ANOVA was conducted at a $\alpha = .05$ level of significance.

The first two-factor ANOVA was used to determine if statistically significant differences in mean STAR IRL scores existed among students receiving reading instruction from teachers in the content areas of communication arts, math, science, and social studies during the Previous Year End to Year 1 Fall timeframe. The results of this test of the main effect for content area were not statistically significant ($F = .371$, $df = 4$, 1342 , $p = .829$). There was not enough evidence to conclude there was a difference in students' mean STAR IRL scores based on teachers' content area.

A second two-factor ANOVA was used to determine if statistically significant differences in mean STAR IRL scores existed among students receiving reading instruction from teachers in the content areas of communication arts, math, science, and social studies during the Year 1 Fall to Year 1 Midpoint timeframe. The results of this test of the main effect for content area were statistically significant ($F = 2.64$, $df = 4$, 1346 , $p = .032$) and indicated there was a difference between at least two mean IRL scores among students receiving reading instruction from teachers in the different content areas of communication arts, math, science, and social studies. Table 17 reports the results of the analysis of the main effect for this ANOVA.

Table 17

Main Effect for Content Area, Year 1 Fall to Year 1 Midpoint Timeframe

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Content X Y1 Fall to Y1 Midpoint	4	5.659	2.640	.032
Error	1346	2.144		

Note: Y = Year

A follow-up post-hoc Tukey's Honestly Significant Difference (HSD) test of the data was conducted to make pairwise comparisons to analyze significant differences between change means. This test allowed for a conservative analysis within multiple comparisons (Salkind, 2008). The Tukey's HSD did not provide evidence of a statistically significant difference among students receiving reading instruction from teachers in the different content areas of communication arts, math, science, and social studies.

A third two-factor ANOVA was used to determine if statistically significant differences in mean STAR IRL scores existed among students receiving reading instruction from teachers in the content areas of communication arts, math, science, and social studies during the Year 1 Midpoint to Year 1 Spring timeframe. The results of this test of the main effect for content area were not statistically significant ($F = .159$, $df = 4, 1351$, $p = .959$). There was not enough evidence to conclude there was a difference in students' mean STAR IRL scores based on teachers' content area.

A fourth two-factor ANOVA was used to determine if statistically significant differences in mean STAR IRL scores existed among students receiving reading

instruction from teachers in the content areas of communication arts, math, science, and social studies during the Year 1 Spring to Year 1 End of year timeframe. The results of this test of the main effect for content area were statistically significant ($F = 7.143$, $df = 4, 1334$, $p = .000$) and indicated there was a difference between at least two mean IRL scores among students receiving reading instruction from teachers in the different content areas of communication arts, math, science, and social studies. Table 18 reports the results of the analysis of the main effect for this ANOVA.

Table 18

Main Effect for Content Area, Year 1 Spring to Year 1 End Timeframe

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Content X Y1 Spring to Y1 End	4	16.047	7.143	.000
Error	1334	2.247		

Note: Y = Year

A follow-up post-hoc Tukey's Honestly Significant Difference (HSD) test of the data was conducted to make pairwise comparisons to analyze significant differences between change means. The Tukey's HSD did not provide evidence of a statistically significant difference among students receiving reading instruction from teachers in the different content areas of communication arts, math, science, and social studies.

A fifth two-factor ANOVA was used to determine if statistically significant differences in mean STAR IRL scores existed among students receiving reading instruction from teachers in the content areas of communication arts, math, science, and

social studies during the Year 1 End to Year 2 Fall timeframe. The results of this test of the main effect for content area were not statistically significant ($F = 1.179$, $df = 4$, 1315, $p = .318$). There was not enough evidence to conclude there was a difference in students' mean STAR IRL scores based on teachers' content area.

A sixth two-factor ANOVA was used to determine if statistically significant differences in mean STAR IRL scores existed among students receiving reading instruction from teachers in the content areas of communication arts, math, science, and social studies during the Year 2 Fall to Year 2 Midpoint timeframe. The results of this test of the main effect for content area were not statistically significant ($F = 1.165$, $df = 4$, 1320, $p = .325$). There was not enough evidence to conclude there was a difference in students' mean STAR IRL scores based on teachers' content area.

A seventh two-factor ANOVA was used to determine if statistically significant differences in mean STAR IRL scores existed among students receiving reading instruction from content area teachers (communication arts, math, science, and social studies) during the Year 2 Midpoint to Year 2 Spring timeframe. The results of this test of the main effect for content area were not statistically significant ($F = 1.307$, $df = 4$, 1339, $p = .265$). There was not enough evidence to conclude there was a difference in students' mean STAR IRL scores based on teachers' content area.

An eighth two-factor ANOVA was used to determine if statistically significant differences in mean STAR IRL scores existed among students receiving reading instruction from content area teachers (communication arts, math, science, and social studies) during the Year 2 Spring to Year 2 End timeframe. The results of this test of the main effect for content area were not statistically significant ($F = 1.789$, $df = 4$, 1350,

$p = .129$). There was not enough evidence to conclude there was a difference in students' mean STAR IRL scores based on teachers' content area.

Table 19 illustrates the sample means for the content areas communication arts, math, science, and social studies for each of the eight two-factor ANOVAs conducted to explore RQ3.

Table 19

Sample STAR IRL Means by Content Area

Test	Timeframe	CA	MATH	SCI	SS
1	Previous Year End to Y1 Fall	.044	.060	.080	.004
2	Y1 Fall to Y1 Midpoint	.341	.222	.429	.316
3	Y1 Midpoint to Y1 Spring	.237	.216	.250	.097
4	Y1 Spring to Y1 End	.373	.373	.266	.510
5	Previous Y1 End to Y2 Fall	.229	.241	.269	.249
6	Y2 Fall to Y2 Midpoint	.327	.341	.225	.099
7	Y2 Midpoint to Y2 Spring	.328	.075	.207	.070
8	Y2 Spring to Y2 End	.307	.220	.147	.356

Note: Y = Year, CA = Communication Arts, Sci = Science, SS = Social Studies

In summary, six of the eight ANOVAs conducted to explore the relationship between students mean IRL scores and teachers' content area demonstrated the main

effect for content area was not statistically significant, indicating there was no difference in mean IRL scores between students taught by communication arts, math, science, and social studies teachers. Follow-up post-hoc Tukey HSD testing of the second and fourth ANOVAs which indicated a difference in student scores related to teachers' content area did not reveal a statistically significance difference among students receiving instruction from teachers in the different content areas of communication arts, math, science, and social studies.

RQ4: To what extent have mean IRL scores on the Renaissance Learning Company's STAR test improved for students based upon the content area teaching assignment of their reading teacher and their own classification as below or at/above grade level proficiency?

H4: There is a difference in the mean IRL change scores students of different content area teachers exhibit if students are below grade level proficiency vs. at/above grade level proficiency.

The 8 two-factor ANOVAs conducted to address RQ3 also addressed RQ4. The two categorical variables used to group the students' mean IRL change scores were teachers' content area (communication arts, math, science, and social studies) and proficiency level (below and at/above grade levels). A two-factor ANOVA can be used to test three hypotheses including a main effect for proficiency level, a main effect for content, and a two-way interaction effect (Content X Proficiency Level). The interaction effect (Content X Proficiency Level) was used to address RQ4, using differences in mean STAR IRL change scores as the dependent variable. Each of the eight ANOVAs was

calculated using mean STAR IRL change score from cohorts 2 and 3, and each ANOVA was conducted at $\alpha = .05$ level of significance.

The interaction effect from the first ANOVA conducted to test the Previous Year End to Year 1 Fall timeframe was used to determine if differences in mean STAR IRL scores among students receiving instruction provided by different content area teachers (communication arts, math, science, social studies) were influenced by the students' proficiency level. The results of the analysis did not indicate statistically significant differences ($F = 2.140$, $df = 4, 1342$, $p = .074$). Table 20 includes the mean, standard deviation (SD), and number (N) of students for the Previous Year End to Year 1 Fall timeframe by content area and proficiency level.

Table 20

STAR Mean IRL Scores by Content Area and Proficiency Level, Previous Year End to Year 1 Fall

Content	Proficiency Level	Mean	SD	N
Communication Arts	At	-.412	1.58	99
	Below	.330	.983	158
	Total	.044	1.30	257
Math	At	-.235	1.76	54
	Below	.169	1.02	146
	Total	.060	1.27	200
Science	At	-.129	2.19	59
	Below	.240	1.00	77
	Total	.080	1.63	136
Social Studies	At	-.470	2.04	75
	Below	.219	1.27	165
	Total	.004	1.58	240

The interaction effect from the second ANOVA conducted to test the Year 1 Fall to Year 1 Midpoint timeframe was used to determine if differences in mean STAR IRL scores among students receiving instruction provided by different content area teachers were influenced by the students' proficiency level. The results of the analysis did yield statistically significant differences ($F = 2.90$, $df = 4$, 1346 , $p = .021$). Table 21 includes the mean, standard deviation (SD), and number (N) of students for the Year 1 Fall to Year 1 Midpoint timeframe.

Table 21

STAR Mean IRL Scores by Content Area and Proficiency Level, Year 1 Fall to Year 1 Midpoint

Content	Proficiency Level	Mean	<i>SD</i>	<i>N</i>
Communication Arts	At	.139	1.73	103
	Below	.469	1.03	163
	Total	.341	1.35	266
Math	At	-.271	1.64	68
	Below	.452	1.23	146
	Total	.222	1.41	214
Science	At	-.107	1.90	46
	Below	.715	1.23	86
	Total	.429	1.54	132
Social Studies	At	-.156	1.79	95
	Below	.661	1.18	130
	Total	.316	1.52	225

Table 21 reports the results of the two-factor (Content Area X Proficiency) ANOVA, indicating a statically significant interaction. The results of this ANOVA indicated that at least two means were different and a follow-up post hoc specified statistically significant differences.

Table 22 reports the analysis of the interaction effect between Content Area and Proficiency, indicating a statistically significant interaction.

Table 22

Results of the Content Area by Proficiency Level Interaction Effect

Source	<i>df</i>	Mean Square	F	<i>p</i>
Y1 Fall to Y1 Midpoint by Content	4	6.217	2.90	.021
Error	1346	2.144		

Note: Y = Year

A follow-up post hoc Tukey's HSD test of the data was conducted to make pairwise comparisons to analyze significant differences between change means. The Tukey's HSD provided evidence of statistically significant differences. Students in the At/Above proficiency level receiving instruction from communication arts teachers (.139) did not experience score decline, while those receiving instruction from math (-.271), science (-.107), and social studies (-.156) teachers exhibited mean decrease. A table included in Appendix F reports the results of the Tukey's HSD for the Year 1 Fall to Year 1 Midpoint timeframe.

The interaction effect from the third ANOVA conducted to test timeframe Year 1 Midpoint to Year 1 Spring was used to determine if differences in mean STAR IRL scores among students receiving instruction provided by different content area teachers were influenced by the students' proficiency level. The results of the analysis did not indicate statistically significant differences ($F = 1.379$, $df = 4, 1351$, $p = .239$). Table 23 includes the mean, standard deviation (SD), and number (N) of students for the Year 1 Midpoint to Year 1 Spring timeframe.

Table 23

STAR Mean IRL Scores by Content Area and Proficiency Level, Year 1 Midpoint to Year 1 Spring

Content	Proficiency Level	Mean	SD	N
Communication Arts	At	-.054	1.76	98
	Below	.410	1.22	165
	Total	.237	1.46	263
Math	At	-.226	1.83	83
	Below	.453	1.33	154
	Total	.216	1.55	237
Science	At	-.163	1.53	56
	Below	.559	1.27	75
	Total	.250	1.43	131
Social Studies	At	-.142	1.76	108
	Below	.351	1.44	103
	Total	.099	1.63	211

The interaction effect from the fourth ANOVA to test timeframe Year 1 Spring to Year 1 End was used to determine if differences in mean STAR IRL scores among students receiving instruction provided by different content area teachers were influenced by the students' proficiency level. The results of the analysis did not indicate statistically significant differences ($F = 1.64$, $df = 4$, 1334 $p = .163$). Table 24 includes the mean, standard deviation (SD), and number (N) of students for the Year 1 Spring to Year 1 End timeframe.

Table 24

STAR Mean IRL Scores by Content Area and Proficiency Level, test Year 1 Spring to Year 1 End

Content	Proficiency Level	Mean	SD	N
Communication Arts	At	-.110	1.84	121
	Below	.805	1.17	135
	Total	.373	1.59	256
Math	At	-.030	1.76	100
	Below	.698	1.33	124
	Total	.373	1.57	224
Science	At	-.197	1.56	59
	Below	.635	1.29	74
	Total	.266	1.47	133
Social Studies	At	.264	1.66	106
	Below	.757	1.48	106
	Total	.510	1.59	212

The interaction effect from the fifth ANOVA to test timeframe Year 1 End to Year 2 Fall was used to determine if differences in mean STAR IRL scores among students receiving instruction provided by different content area teachers were influenced by the students' proficiency level. The results of the analysis did indicate statistically significant differences ($F = 3.317$, $df = 4$, 1315, $p = .019$). Table 25 includes the mean, standard deviation (SD), and number (N) of students for the Year 1 End of Year to Year 2 Fall timeframe.

Table 25

STAR Mean IRL Scores by Content Area and Proficiency Level, Year 1 End of Year to Year 2 Fall

Content	Proficiency Level	Mean	SD	N
Communication Arts	At	-.147	1.74	122
	Below	.473	1.42	188
	Total	.229	1.58	310
Math	At	-.735	1.74	113
	Below	.697	1.63	242
	Total	.241	1.79	355
Science	At	-.453	1.92	133
	Below	.802	1.57	167
	Total	.269	1.64	315
Social Studies	At	-.453	1.92	133
	Below	.693	1.48	210
	Total	.249	1.75	343

Table 25 reports the results of the two-factor (Content Area X Proficiency) ANOVA, indicating a statically significant interaction. The results of this ANOVA indicated that at least two means were different and a follow-up post hoc specified statistically significant differences.

Table 26 reports the analysis of the interaction effect between Content Area and Proficiency, indicating a statistically significant interaction.

Table 26

Results of the Content Area by Proficiency Level Interaction Effect

Source	<i>df</i>	Mean Square	F	<i>p</i>
Y1 End of Year to Y2 Fall by Content	3	8.587	3.317	.019
Error	1315	2.589		

Note: Y = Year

A follow-up post hoc Tukey's HSD test of the data was conducted to make pairwise comparisons to analyze significant differences between change means. The Tukey's HSD provided evidence of statistically significant differences. A table included in Appendix G reports the results of the Tukey's HSD for the Year 1 End of Year to Year 2 Fall timeframe. Students in the At/Above proficiency level receiving reading instruction by math teachers exhibited greater score decline (-.735) than those receiving instruction from communication arts teachers (-.177).

The interaction effect from the sixth ANOVA to test the Year 2 Fall to Year 2 Midpoint timeframe was used to determine if differences in mean STAR IRL scores among students receiving instruction provided by different content area teachers were influenced by the students' proficiency level. The results of the analysis did not yield statistically significant differences among mean IRL scores ($F = 1.724$, $df = 3, 1320$, $p = .160$). Table 27 includes the mean, standard deviation (SD), and number (N) of students for the Year 2 Fall to Year 2 Midpoint timeframe.

Table 27

STAR Mean IRL Scores by Content Area and Proficiency Level, Year 2 Fall to Year 2

Midpoint

Content	Proficiency Level	Mean	SD	N
Communication Arts	At	-.056	1.41	123
	Below	.565	1.33	198
	Total	.327	1.40	321
Math	At	-.298	1.73	127
	Below	.716	1.43	216
	Total	.341	1.62	343
Science	At	-.115	1.33	172
	Below	.603	1.47	154
	Total	.225	1.44	326
Social Studies	At	-.472	1.73	155
	Below	.583	1.30	183
	Total	.099	1.60	338

The interaction effect from the seventh ANOVA to test timeframe Year 2 Midpoint to Year 2 Spring was used to determine if differences in mean STAR IRL scores among students receiving instruction provided by different content area teachers were influenced by the students' proficiency level. The results of the analysis did yield statistically significant differences among mean IRL scores ($F = 3.56$, $df = 3, 1339$, $p = .014$). Table 28 includes the mean, standard deviation (SD), and number (N) of students for the Year 2 Midpoint to Year 2 Spring timeframe.

Table 28

STAR Mean IRL Scores by Content Area and Proficiency Level, Year 2 Midpoint to Year 2 Spring

Content	Proficiency Level	Mean	SD	N
Communication Arts	At	-.194	1.43	107
	Below	.604	1.41	202
	Total	.328	1.47	309
Math	At	-.492	1.59	157
	Below	.546	4.47	189
	Total	.075	1.61	346
Science	At	.006	1.16	152
	Below	.404	1.24	150
	Total	.204	1.21	302
Social Studies	At	-.452	1.61	199
	Below	.613	1.70	191
	Total	.070	1.74	390

Table 28 reports the results of the two-factor (Content Area X Proficiency) ANOVA, indicating a statically significant interaction. The results of this ANOVA indicated that at least two means were different and a follow-up post hoc specified statistically significant differences.

Table 29 reports the analysis of the interaction effect between Content Area and Proficiency, indicating a statistically significant interaction.

Table 29

Results of the Content Area by Proficiency Level Interaction Effect

Source	<i>df</i>	Mean Square	F	<i>p</i>
Y2 Midpoint to Y2 Spring by Content	3	7.726	3.553	.014
Error	1339	2.174		

Note: Y = Year

A follow-up post hoc Tukey's HSD test of the data was conducted to make pairwise comparisons to analyze significant differences between change means. The Tukey's HSD did not provide evidence of a statistically significant difference among students receiving reading instruction from teachers in the different content areas of communication arts, math, science, and social studies. A table included in Appendix H reports the results of the Tukey's HSD for the Year 2 Midpoint to Year 2 Spring timeframe.

The interaction effect from the eighth ANOVA was used to determine the effect of the interaction between reading instruction provided by different content area teachers (communication arts, math, science, social studies) and classification in different proficiency levels (below or at/above grade level). Table 30 includes the mean, standard deviation (SD), and number (N) of students for the Year 2 Spring to Year 2 End timeframe.

Table 30

STAR Mean IRL Scores by Content Area and Proficiency Level, Year 2 Spring to Year 2

End

Content	Proficiency Level	Mean	<i>SD</i>	<i>N</i>
Communication Arts	At	-.349	1.55	123
	Below	.668	1.47	223
	Total	.307	1.58	346
Math	At	-.162	1.57	167
	Below	.586	1.40	174
	Total	.220	1.53	341
Science	At	-.125	1.26	160
	Below	.455	1.54	138
	Total	.144	1.42	298
Social Studies	At	-.219	1.61	207
	Below	1.07	1.76	166
	Total	.356	1.80	373

Table 31 reports the analysis of the interaction effect between Content Area and Proficiency. The results of this interaction effect yielded statistically significant results for this timeframe.

Table 31

Results of the Content Area by Proficiency Level Interaction Effect, Year 2 Spring to Year 2 End Timeframe

Source	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>
Y2 Spring to Y2 End	3	8.161	3.497	.015
Error	1350	2.334		

Note: Y = Year

A follow-up post hoc Tukey's HSD test of the data was conducted to make pairwise comparisons to analyze significant differences between change means. The Tukey's HSD did not provide evidence of a statistically significant difference among students receiving reading instruction from teachers in the different content areas of communication arts, math, science, and social studies. A table included in Appendix G reports the results of the Tukey's HSD for the Year 2 Spring to Year 2 End timeframe.

In summary, four of the 8 two-factor ANOVAs conducted to explore the effect of the relationship between teachers' content area and student proficiency levels demonstrated the interaction effect for content area was statistically significant. Follow-up post-hoc Tukey's HSD testing of these ANOVAs only revealed statistically significant interaction for the Year 1 End to Year 2 Fall timeframe. No other post-hocs yielded statistically significant interaction effects. Additionally, four of the eight ANOVAs exploring this relationship did not indicate significance for this interaction.

Summary

In this chapter, frequency data regarding the number of students representing each school and each of the three cohorts included in the study were provided. This

chapter also presented the results of the statistical analyses using t test and ANOVA hypothesis testing. Results of the analyses were presented as well. Chapter five summarizes the study, explores the findings, and discusses their connection to the literature. Implications for action, recommendations for future research, and conclusions are also included.

Chapter Five

Interpretation and Recommendations

One positive aspect of the No Child Left Behind Act has been an increased awareness regarding the importance of providing evidence-based reading instruction at every level of public education (Allington, 2006b; Luke & Woods, 2007). This is particularly true regarding the long-overlooked area of adolescent literacy. As society turned its eye toward academic achievement in this area, it was essential that educators continually monitored instructional effectiveness and addressed ongoing student needs (Rissman, Miller, & Torgesen, 2009). The present study focused on evaluating the effectiveness of a reading instruction intervention, the Success for All (SFA) Foundation's Reading Edge program, utilized in three middle schools in the Independence Public School District. Chapter one introduced the background, purpose, and significance of the present study. Chapter two presented a review of literature that explained the growing focus on adolescent literacy, relevant studies about the SFA Foundation's middle school Reading Edge curriculum package, and effective adolescent literacy instruction. Chapter three detailed the methodology of the study, and chapter four presented the results of hypothesis testing related to the research questions. This chapter presents a brief review of the problem, purpose, research questions, methodology, and major findings of the study. Additionally, connections between the present study and relevant literature describing SFA studies, scripted curriculum, and teaching out-of-field are presented with implications for action and recommendations for future research.

Study Summary

The first section of this chapter provides a condensed summary of the present study. To begin, the summary contains an overview of the status of adolescent literacy as a national concern. Next, the second section explains the purpose of the study. The third section reviews methodology used in the study, and the fourth presents major findings of the study.

Overview of the Problem. Despite attempts at education reform and an increased public focus on reading achievement, students in the United States continued to struggle to meet grade-level proficiency standards in reading (National Center for Educational Statistics, 2011). Middle-level education was no exception to this problem. According to the National Center for Educational Statistics (NCES) long-term trend report, National Assessment of Educational Progress (NAEP) scores for 13-year olds have remained stagnant since 1994 (Nation's Report Card, 2011). The scores have consistently fallen within the basic proficiency level, indicating only partial mastery of the knowledge and skills considered fundamental for a designated grade (NCES, 2011). Consequently, schools have adopted a variety of instructional approaches intended to improve reading achievement.

Purpose Statement and Research Questions. The purpose of this study was to analyze the effect of implementation of the Reading Edge program, a scripted curriculum offered by the SFA Foundation, on reading achievement in three, non-Title I middle schools. The impact of this instructional approach was measured by the Renaissance Learning Company's Standardized Test for the Assessment of Reading (STAR) Instructional Reading Level (IRL). This study explored the extent to which there was a

statistically significant difference in mean STAR IRL score changes for students categorized at proficiency levels of either below grade level or at/above grade level. A second purpose of the study was to explore the relationship between mean IRL scores of below or at/above grade level proficiency students and their instructors' content teaching areas, determining whether the knowledge and pedagogy of teachers in a particular content area correlated with greater academic gains for students.

Review of the Methodology. The design of this study was quantitative, non-experimental, and causal-comparative in nature. According to Johnsen and Christen (2008), causal-comparative research exists when “there is one categorical independent variable and one quantitative dependent variable” (p. 360). In each research question in the study, students' mean Instructional Reading Level (IRL) scores on quarterly Renaissance Learning Company's STAR assessments were compared, acting as a quantitative dependent variable. The sample size was 1,470 students enrolled in three, non-Title I Independence School District (ISD) middle schools between 2007 and 2010. All students in the sample completed nine STAR assessments during their designated window of program participation. Once the data were accessed from archival records, compiled into a Microsoft Excel 2010 spreadsheet, and coded for anonymity, the researcher utilized the spreadsheet for analysis of Hypotheses One (H1) and Hypothesis Two (H2), conducting 8 one-sample *t* tests for each proficiency level to determine significance. The sample data were also input into the IBM Statistical Package for the Social Sciences (SPSS) Faculty Pack 19.0 for analysis of Hypothesis Three (H3) and Four (H4). A series of eight different two-factor ANOVAs were utilized to analyze data for H3, evaluating the main effect for content area. The same series of 8 two-factor

ANOVAs were also conducted for H4, but were utilized to evaluate the interaction effect of content area by proficiency level. Tukey's HSD post hoc analyses were used to study significant interaction effects within the ANOVAs. Additional dependent variables included student proficiency levels and teachers' content area.

Major Findings. The researcher investigated mean IRL score change for students enrolled in the ISD who received instruction via the SFA Foundation's Reading Edge scripted curriculum package. This investigation included only students for whom nine STAR test records were available within a designated timeframe, eliminating transiency as an influencing factor when considering achievement.

The data analysis yielded statistically significant results in each of the eight quarterly *t* tests conducted for the below grade-level proficiency group, with each of the eight tests revealing a positive mean IRL score change on regular administration of the STAR. This indicates that students in this proficiency level did benefit from the time specifically designated for reading instruction.

Additionally, data analysis yielded statistically significant results in each of the eight *t* tests conducted for students in the at/above grade level proficiency designation. However, in six of the eight tests for this group, the mean IRL score decreased, indicating these students did not demonstrate improvement on regular administration of the STAR but instead often experienced IRL score decline.

In analyzing ANOVA results for testing related to H3, there was no consistent significant difference for students taught by instructors categorized in the different content areas of communication arts, math, science, and social studies. These ANOVA

results support SFA's suggestion that all certified staff can achieve equal results utilizing the scripted curriculum package.

The ANOVA results for H4 also revealed few statistically significant differences for below or at/above grade level proficiency students of communication arts, math, science, or social studies teachers, indicating that no particular content area's teachers tended to be more effective in providing reading instruction to either proficiency group.

Findings Related to the Literature

Multiple studies have been published about the elementary SFA reading program, but few have been conducted about the middle school program, the Reading Edge. One such study including site-level implementation data analysis was conducted by Chamberlain, Daniels, Madden, and Slavin (2007), researchers affiliated with the SFA Foundation or with Johns Hopkins University. Testing at the end of this one-year study comparing achievement growth in 405 experimental and control group sixth graders found modest effect sizes (p. 13), similar to the mixed results of the present study. The data for the sixth grade participants in the Chamberlain, et al. (2007) study were examined for the entire group rather than divided into proficiency levels for analysis. At the end of their year of program participation, the experiment and control groups both took the Gates MacGinitie Reading Test, allowing researchers to compare their academic progress. Testing of students in the experiment group (N = 203) revealed a moderate effect size (ES = +.14) on the Gates Total score, indicating that implementation of the instructional program was somewhat effective in positively influencing student achievement.

In the present study, testing utilizing the Renaissance Learning Company's STAR consistently revealed positive, statistically significant mean IRL score changes for below grade level proficiency students (see Table 15). However, students who ranked at/above grade level proficiency exhibited statistically significant mean IRL decreases on six of the eight *t* tests for their group (see Table 16). These mixed results indicate that the instructional program was somewhat effective in influencing overall student academic achievement, though it appeared to have greater positive impact for below grade-level proficiency students.

In contrast to the present study, Chamberlain, Daniels, Madden, and Slavin (2007) examined elements of instructional fidelity to determine their effect on achievement data, concluding that a greater effect size might be obtained after teachers pass the initial-implementation phase of program usage. However, they did not consider overall program fidelity (i.e. moving students to different classes on a quarterly basis or having a full-time program facilitator) or peripheral structural elements (tutoring, utilizing a Success Team), because there were limitations imposed by the experimental setting. Some of the imposed limitations included the fact that the schools were not able to follow the Foundation's recommendation to change students' classes on a quarterly basis, nor did the schools in their study have a full-time program facilitator to provide professional or organizational support. The three schools participating in the present study did assess students on a quarterly basis and changed classes based upon assessment data, in accordance with Foundation program recommendations. Each of the three schools in the present study also had a full-time facilitator on staff to provide professional development, instructional assistance, and ongoing organizational support. However, none of the three

schools opted to provide formal reading tutoring related to SFA program use, and none implemented a Success Team for program leadership. Additionally, the present study did not account for elements of instructional fidelity in using the scripted materials.

Another study, also conducted by researchers affiliated with the SFA Foundation or Johns Hopkins University, examined state reading test achievement data from 2001 to 2004 in seven control and seven experimental schools located in six states (Daniels, Madden, & Slavin, 2004). In this analysis of achievement data obtained from state department of education websites, researchers found the SFA schools made significant gains in the percent of students passing state reading tests and made greater gains than each of the control schools. Daniels, Madden, and Slavin (2004) acknowledged that it was difficult to compare achievement data from state-to-state without analyzing how reading achievement was measured on each of these instruments (p.8). However, they explained it was worth noting that the SFA schools exhibited greater gains than the paired controls. It was also noteworthy that the study took place over an extended timeframe, similar to that of the present study. The present study did not analyze state achievement data as a quantitative variable, though such measurement is relevant in the era of No Child Left Behind. The measure used in the present study, mean Instructional Reading Level (IRL) scores on the Renaissance Place Standardized Test for the Assessment of Reading (STAR), did indicate consistent, statistically significant change for students categorized below grade level proficiency during implementation of the Reading Edge program. In contrast, students in the at/above grade level proficiency designation tended to demonstrate mean IRL score decline on the same testing instrument.

Finally, a meta-analysis by Sparks (2004) found that evidence in six reading instruction research studies pointed toward a positive effect of qualified, certificated teachers providing reading instruction (Sparks, 2004). The present study examined mean STAR IRL score changes among students classified as below or at/above grade proficiency levels and receiving reading instruction from teachers typically certified to teach classes of communication arts, math, science, and social studies. Results from the quantitative data analysis were mixed, and no qualitative data regarding instructional fidelity, classroom environment, teacher efficacy, or other fields of certification were included. Consequently, it was not possible to ascertain whether the mean IRL score changes that may have been caused by variance of teacher effectiveness or by certification as a reading instructor. However, the field of literature indicated these factors were likely influential (Johnson, 2005; Moats, 1999; Sparks 2004).

Conclusions

As discussed in chapter one, school and district leaders face challenging decisions about literacy instructional programming in the face of sometimes unclear or contradictory research. The findings from this study have implications for numerous stakeholders ranging from school-level leaders to those entrusted with forming state and national policies. The following section delineates implications for action for a variety of parties.

Implications for Action. In the present analysis of mean IRL score changes, students classified as below grade-level proficiency experienced significant and consistent positive mean IRL score change, while those at/above grade level proficiency typically exhibited mean IRL score decline. Consequently, implications for the SFA

Foundation emerged as possible courses of future action. As represented in the SFA Foundation's (2004) *The Reading Edge: Overview* (2004) booklet, the written format of the scripted curriculum changes very little within the SFA levels 4-8+ range. This is the packaged curriculum area designated for students ranging from slightly below to the at/above grade level proficiency group. This could be a factor in the decreasing mean IRL score changes among students in the present study. The Foundation may want to institute a series of professional development sessions addressing reading development and pedagogy, focusing this training on providing instruction via a scripted curriculum package to students who are deemed to be at the appropriate proficiency level. Additionally, the SFA Foundation's scripted curriculum package materials for levels 6-8+ is intended for use with students who are at or above grade level proficiency, depending upon their grade in school. The repeated mean IRL score decrease may indicate these materials are not fully suitable for use with students at this proficiency level. The Foundation may want to analyze the effectiveness of current materials with this particular proficiency group.

Having consulted a district program facilitator for input about program implementation within the ISD, the researcher predicted the decreasing mean IRL scores may also have been due to larger class sizes at the higher-proficiency instructional levels. Typical class sizes at these levels within the program tended to be larger than the lower, below-grade proficiency classes, and this may have decreased instructors' abilities to interact with students to meet their individual needs. The SFA Foundation may want to explore making recommendations about class sizes to schools using their scripted curriculum materials (L. McGee, personal communication, February 18, 2012).

The review of literature indicated reading pedagogy training among secondary-certified instructional staff to be an area of weakness (Meyer, 2009; Sparks, 2004). Findings of the present study somewhat refuted the literature, as below-grade level proficiency students receiving reading instruction from communication arts, math, science, and social studies teachers utilizing the SFA Foundation's scripted Reading Edge curriculum package exhibited positive, statistically significant mean IRL score change on regular STAR test administration.

Additionally, findings of the present study revealed communication arts, math, science, and social studies certified staff consistently obtained statistically comparable student achievement data. However, when considering the similar ANOVA results in conjunction with the declining mean IRL scores for at/above grade level proficiency on the aforementioned *t* tests, the similarities may indicate that as a group, many middle-level educators enter the classroom equally unprepared to teach the craft of reading. Together, these factors have implications for school leaders and professional development staff. As districts strive to improve the quality of education they provide, and as requirements set forth by No Child Left Behind (2002) continue to increase, it is imperative that districts provide ongoing, effective, multi-faceted literacy professional development on a regular basis for all teachers.

Finally, findings from the present study have a variety of implications for policy makers at the state and national level. These include:

- Revise certification standards to require secondary certified teachers to have more training in literacy instruction prior to entering the field.

- Require more frequent professional development in literacy for secondary instructors already certified and in the field.
- Evaluate whether scripted curriculum programs that utilize all certified staff, creating a situation in which many provide instruction in an out-of-field area, should be exempt from highly-qualified coding requirements at the state Department of Elementary and Secondary Education level.

Recommendations for Future Research. Findings from the current study warrant further research in several areas.

- First, replicate the present study, including an analysis of instructional fidelity to the scripted curriculum guidelines. Such a study would also benefit from a line of inquiry addressing state testing achievement data.
- Second, conduct further long-term research on the Reading Edge program with all of the components and structural supports in place. Such an analysis should constitute an examination of recommended components including regular assessment, moving students to different instructional levels, a full-time program coordinator, etc. An analysis of structural supports should include tutoring and the use of a leadership (Solutions) team. It may be beneficial to know which components are most effective in improving achievement.
- Third, future research should study Reading Edge program implementation in a way that compares those instructors who have higher/stronger/more intense levels of reading pedagogy preparation, such as that typically provided in an elementary certification program, versus those who are traditionally secondary

certified without much reading pedagogy training. Ultimately, if teachers are assigned out-of-field teaching responsibilities including reading instruction, it is important to understand how their preparation via certification through teacher training programs affects ongoing needs for professional development, program implementation support, and other factors.

- Fourth, future research should study Reading Edge program implementation in a way that analyzes its effect on achievement on English language arts-reading assessments. In the state of Missouri, where the present study was conducted, this would include an analysis of Missouri Assessment Program (MAP) scores.
- Finally, future researchers should evaluate the impact of No Child Left Behind (2002) requirements on scripted program use, and explore whether this type of packaged curriculum improves long-term academic achievement.

Concluding Remarks. This study examined 2007 to 2010 achievement data via the STAR IRL scores in three, non-Title I middle schools during implementation of the SFA Foundation's Reading Edge scripted curriculum instructional program. Analysis of data revealed significant change for students deemed below grade level proficiency and also revealed a pattern of significant score decline for students categorized at/above grade level proficiency. The data were further analyzed to determine if there was a difference in achievement of students receiving instruction from communication arts, math, science, and social studies teachers. Data indicated no consistently significant differences in achievement for students of either proficiency level taught by teachers of communication

arts, math, science, or social studies when utilizing the Reading Edge, a scripted curriculum material package.

Reading and literacy achievement continues to be an issue of national concern (Biancarosa & Snow, 2004; National Center for Educational Statistics, 2011). Schools and districts must find instructional methods and programs to support struggling readers and stimulate further growth among those already considered proficient. Such programs should address the unique needs of students as individual learners, build upon their capacity, and address expectations set forth by state guidelines. By enabling students to meet reading proficiency goals, we not only help them gain greater academic success, but we also empower them to enter the future with a vital tool for success as a citizen of the 21st Century- the ability to read.

References

- African American Voices in Congress. (2012). *Educational Reform and the CBC: Spotlight on the Elementary and Secondary Education Act (ESEA)*. Retrieved from: <http://www.avoiceline.org/edpol/index.html>
- Alexander, P., & Jetton, T. (2000). Learning from text: A multidimensional developmental perspective. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, p. 285-310). Mahaw, NJ: Erlbaum.
- Alliance for Excellent Education. (2011, December). Adolescent Literacy. *Fact Sheet*. Retrieved from <http://www.all4ed.org/files/AdolescentLiteracyFactSheet.pdf>
- Allington, R. (2002). You can't learn from books you can't read. *Educational Leadership*, 60(3), 16-19.
- Allington, R. (2006a). Troubling times: A short historical perspective. In R.L. Allington (Ed.), *Big brother and the national reading curriculum* (pp. 3-46). Portsmouth, NH: Heinemann.
- Allington, R. (2006b). *What really matters for struggling readers*. Boston, MA: Pearson Education, Inc.
- American Institutes for Research. (2005). *CSRQ Center report on elementary school comprehensive school reform models*. Washington, DC: The Comprehensive School Reform Quality Center American Institutes for Research.

- Atkinson, C. (1998). An analysis of the impact of Success for All on reading, attendance, and academic self-efficacy with at-risk elementary school students (Doctoral dissertation, Virginia Polytechnic Institute and State University). Retrieved from <http://scholar.lib.vt.edu/theses/available/etd-32098-1546/unrestricted/DIS2.PDF>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215.
- Benicoff-Nan, L. (2002). A correlation of computer adaptive, norm referenced, and criterion referenced achievement tests in elementary reading. (Unpublished doctoral dissertation). The Boyer Graduate School of Education, Santa Ana, CA.
- Bergman, I., & Biancarosa, G. (2005). *Reading to achieve: A governor's guide to adolescent literacy*. Washington, DC: National Governors Association Center for Best Practices.
- Biancarosa, C., & Snow, C.E. (2004). *Reading next- A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York (2nd ed.)*. Washington, DC: Alliance for Excellent Education.
- Blankstein, A. (2004). *Failure is not an option*. Thousand Oaks, CA: Corwin Press.
- Blanton, W., Wood, K., & Taylor, D.B. (2007). Rethinking middle school reading instruction: A basic literacy activity. *Reading Psychology*, 28(1), 75-95.
- Blendinger, J., & Wells, L. R. (2001). *Investigating the Success for All reading program*. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Little Rock, Arkansas.

- Borman, G., Slavin, R., Cheung, A., Chamberlain, A., & Madden, A. (2004). *Success for All: Preliminary first-year results from the national randomized field trial*. Baltimore, MD: Success for All Foundation.
- Borman, G., Slavin, R., Cheung, A., Chamberlain, A., Madden, N., & Chambers, B. (2005). Success for All: First-year results from the national randomized field trial. *Educational Evaluation and Policy Analysis*, 27(1), 1-22.
- Borman, G., Slavin, R., Cheung, A., Chamberlain, A., Madden, N., & Chambers, B. (2007). Final reading outcomes of the national randomized field trial of Success for All. *American Educational Research Journal*, 44(3), 701-731. doi: 10.3102/0002831207306743
- Brown, C. (2002). Opportunities and accountability to leave no child behind in the middle grades: An examination of the No Child Left Behind Act of 2001. New York, NY: Edna McConnell Clark Foundation.
- Camilli, G., & Wolfe, P. (2004). Research on reading: A cautionary tale. *What Research Says About Reading*, 61(6), 26-29.
- Cantrell, S., & Hughes, H. (2008). Teacher efficacy and content literacy implementation: An exploration of the effects of extended professional development with coaching. *Journal of Literacy Research*, 40(1), 95-127. doi: 10.1080/10862960802070442
- Carnegie Council on Adolescent Development. (1989). *Turning points: Preparing American youth for the 21st century*. New York, NY: Carnegie Corporation.
- Carnine, D.W., Silbert, J., Kame'enui, E. J., & Tarver, S. G. (2004). *Direct instruction in reading (4th ed.)*. Upper Saddle River, NJ: Pearson-Prentice Hall.

- Center for Collaborative Education. (n.d.). Transforming middle schools. *Turning points*. Retrieved from <http://www.turningpts.org/index.html>.
- Chall, J. S. (1995). *The stages of reading development (2nd ed.)*. New York, NY: Wadsworth.
- Chamberlain, A., Daniels, C., Madden, N. A., & Slavin, R. E. (2007). A randomized evaluation of the Success for All middle school reading program. *Middle Grades Research Journal*, 2(1), 1-21.
- Christenbury, L., Bomer, R., & Smagorinsky, P. (2009). *Handbook of adolescent literacy research*. New York: Guilford Press.
- Clarke, C. (2002). *Reading Intervention Programs: A best-practices comparison of ten middle school practitioners in San Diego County, California* (Doctoral dissertation.) Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3042712)
- Coles, G. (2001). Reading taught to the tune of a scientific hickory stick. *Phi Delta Kappan*, 83(3), 204-212.
- Commeyras, M. (2007). Scripted reading instruction? What's a teacher educator to do? *Phi Delta Kappan*, 88 (5), 404-407.
- Corum, J., Kepler, D., Mattson, T., & Okerstrom, J. (2007). *An examination of secondary literacy practices: An evaluative tool for school leaders to improve literacy instruction and student achievement* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3280191)
- Daniels, C., Madden, N., & Slavin, R. E. (2004). *The Success for All middle school: Third year evaluation report*. Retrieved from ERIC database. (ED 483813)

- Darling-Hammond, L., & Youngs, P. (2002). Defining “highly qualified teachers”: What does “scientifically-based research” actually tell us? *Educational Researcher*, 31(9), 13-25.
- Darling-Hammond, L., Holtzman, D., Gatlin, S., & Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Educational Policy Analysis Archives*, 13(1), Retrieved from <http://www.redalyc.org/src/inicio/ArtPdfRed.jsp?iCve=275020513042>
- Dennis, D.V. (2007). *Putting the horse before the cart: Utilizing what assessment data reveal about struggling young adolescent readers to inform policy area instruction* (Doctoral dissertation). Retrieved from Proquest Dissertations and Theses database. (UMI No. 3286915)
- Denton, J., & Laciana, L. (1982). *Quantity of professional education coursework linked with process measures of student teaching*. Retrieved from ERIC database. (ED 219355)
- Deschler, D., Palincsar, A., Biancarosa, G., & Nair, M. (2007). *Informed choices for struggling adolescent readers: A research-based guide to instructional programs and practices*. New York, NY: International Reading Association.
- Dianda, M., & Flaherty, J. (1995). *Effects of Success for All on the reading achievement of first graders in California bilingual programs*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Dillon, S. (2011, August 8). Overriding a key education law. *New York Times*, pp. A12.
- DuBay, W. H. (2004). *The principles of readability*. Retrieved from <http://www.nald.ca/fulltext/readab/readab.pdf>

- Duncan-Owens, D. (2009). Scripted reading programs: Fishing for success. *Principal*, 88(3), 26-29.
- Ede, A. (2006). Scripted curriculum: Is it a prescription for success? *Childhood Education*, 83(1), 29-32.
- Ehri, L., & Stahl, S. (2001). Beyond the smoke and mirrors: Putting out the fire. *Phi Delta Kappan*, 83(1), 17-20.
- Federal Register. (2008). *Notices*, 73 (69), 1-2.
- Fisher, D., & Ivey, G. (2006). Evaluating the interventions for struggling adolescent readers. *Journal of Adolescent and Adult Literacy*, 50(3), 180-189.
- Fives, H. (2003). *What is teacher efficacy and how does it relate to teachers' knowledge?*
Paper presented at the American Educational Research Association Annual Conference, Chicago, IL.
- Gallagher, K. (2009). *Readicide*. Portland, ME: Stenhouse Publishers.
- Garan, E. (2001). What does the report of the National Reading Panel really teach us about teaching phonics? *Language Arts*, 79(1), 61-71.
- Garan, E. (2002). *Resisting reading mandates: How to triumph with the truth*.
Portsmouth, NH: Heinemann.
- Garan, E. (2005). Murder your Darlings: A scientific response to the voice of evidence in reading research. *Phi Delta Kappan*, 86(6), 438-443.
- Goldhaber, D., & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation Policy and Analysis*, 22(2), 129-145.

- Hawk, P., Coble, C., & Swanson, M. (1985). Certification: It does matter. *Journal of Teacher Education*, 36(3), 13-15.
- Heller, R., & Greenleaf, C. (2007). *Literacy instruction in the content areas: Getting to know the core of middle and high school improvement*. Washington, DC: Alliance for Excellent Education.
- Hill, J. (2011). *Education and certification qualifications of departmentalized public high school teachers of core subjects: Evidence from the 2007-2008 schools and staffing survey*. Washington, DC: National Center for Educational Statistics.
- Hinson, J. (2010). *Inspiring greatness*. Independence, MO: Independence School District.
- Hurley, E., Chamberlain, A., & Slavin, R. (2001). Effects of Success for All on TAAS reading scores. *Phi Delta Kappan*, 82(10), 750-756.
- Ingersoll, R.M. (2003). *Out-of-field teaching and the limits of teacher policy*. (CTP Report R-03-5). Retrieved from http://www.archives.nysed.gov/edpolicy/research/res_essay_intro_fedrole_states.shtml
- International Reading Association. (2000). *Excellent reading teachers. A position statement of the International Reading Association*. Newark, DE: International Reading Association.
- Institute of Educational Sciences. (2012, February 10). National Assessment of Educational Progress (NAEP). National Center for Education Statistics. Retrieved from <http://nces.ed.gov/nationsreportcard/ltr/>
- Ivey, G., & Broaddus, K. (2000). Tailoring the fit: Reading instruction and middle school readers. *The Reading Teacher*, 54(1), 68-78.

- Jackson, A.W., & Davis, G.A.(2000). *Turning points 2000: Educating adolescents in the 21st Century*. New York, NY: Teacher's College Press.
- Jennings, J., & Renter, D. (2006). How public schools are impacted by 'No Child Left Behind'. *Education Digest: Essential Readings for Quick Review*, 72(4), 4-9.
- Jetton, T. L., & Alexander, P. A. (2004). Domains, teaching, and literacy. In T. L. Jetton & J.A. Dole (Eds.), *Adolescent literacy research and practice* (pp. 15–39). New York, NY: Guilford Press.
- Joftus, S., & Maddox-Dolan, B. (2003). *Left out and left behind: NCLB and the American high school*. Washington, DC: Alliance for Excellent Education.
- Johnson, T. (2005). *An examination of the relationship between the type of teacher certification of middle school teacher sand student math, reading, and writing achievement* (Doctoral dissertation). Retrieved from Proquest Dissertations and Theses database. (UMI No.31640800)
- Johnson, R., and Christensen, L. (2008). *Educational research quantitative, qualitative, and mixed approaches*. Los Angeles, CA: Sage Publications.
- Joyce, B. (1999).The greatest literacy problem and Success for All. *Phi Delta Kappan*, 81(2), 129-131.
- Kamil, M. (2003). *Adolescents and literacy: Reading for the 21st century*. Washington, DC: Alliance for Excellent Education.
- Klingner, J., Cramer, E., & Harry, B. (2006). Challenges in the implementation of Success for All in four high-need urban schools. *The Elementary School Journal*, 106(4), 333-349.

- Laczko-Kerr, I., & Berliner, D.C. (2002). The effectiveness of “Teach for America” and other under-certified teachers on student academic achievement: A case of harmful policy. *Education Policy Analysis Archives, 10* (37), 1-56.
- Learning Point. (2004). *A closer look at the five essential components of effective reading instruction: A review of scientifically based reading research for teachers*. Naperville, IL: Learning Points Associates.
- Luke, A., & Woods, A. (2007). Learning lessons: What No Child Left Behind can teach us about literacy, testing, and accountability. *Professional Magazine, 22*, 1-8. Retrieved from http://www.acsa.edu.au/pages/images/learning%20lessons_from_no_child_left_behind.pdf
- Lunenburg, F., & Irby, B. (2008). *Writing a successful thesis or dissertation*. Thousand Oaks, CA: Corwin Press.
- Mallette, M., Henk, W., Waggoner, J., & Delaney, C. (2005). What matters most? A survey of accomplished middle-level educators’ beliefs and values about literacy. *Action in Teacher Education, 27*(2), 33-42.
- Manzo, A., Manzo, U., & Thomas, M. (2005). *Content area literacy: Strategic teaching for strategic learning*. Hoboken, NJ: Wiley & Sons.
- McEwan, E. (2007). *Raising reading achievement in middle and high schools*. Thousand Oaks, CA: Corwin Press, Inc.
- Medwell, J., Wray, D., Poulson, L., & Fox, R. (1998). *Effective teachers of literacy*. Retrieved from <http://www.leeds.ac.uk.educol/documents/000000829.htm>

- Meyer, C. (2009). *An assessment of middle and high school content-area teachers' knowledge of adolescent literacy* (Doctoral dissertation). Retrieved from Proquest Dissertations and Theses database. (UMI No. 3373036)
- Missouri Department of Elementary and Secondary Education. (2011a). *2010 Missouri kids count data book online*. Jefferson City, MO: Author. Retrieved from http://mcdc2.missouri.edu/pub/webrepts/kidscnt/outcomes_facts/29095.html
- Missouri Department of Elementary and Secondary Education. (2011b). *Comprehensive data system: Independence 30 (048077)*. Jefferson City, MO: Author. Retrieved from: <http://mcds.dese.mo.gov/guidedinquiry/District%20and%20Building%20Student%20Indicators/Building%20Demographic%20Data.aspx?rp:District=048077&rp:SchoolYear=2011&rp:SchoolYear=2010&rp:SchoolYear=2009&rp:SchoolYear=2008>
- Missouri Department of Elementary and Secondary Education. (2012). *Compendium of Missouri certification requirements index*. Jefferson City, MO: Author. Retrieved from <http://www.dese.mo.gov/schoollaw/rulesregs/EducCertManual/Index.htm>
- Moats, L. (1999). *Teaching reading is rocket science: What expert teachers of reading should know and be able to do*. Washington, D.C.: National Institute of Child Health and Human Development.
- Moore, D., Bean, T., Birdyshaw, D., & Rycik, J. (1999). *Adolescent literacy: A position statement for the commission on the adolescent literacy of the International Reading Association*. Newark, DE: International Reading Association.

- National Center for Educational Statistics. (2011). Nation's Report Card. Retrieved from http://nationsreportcard.gov/ltt_2008/ltt0009.asp?subtab_id=Tab_2&tab_id=tab1#chart
- National Council of Teachers of English. (2006, April). *NCTE Principles of adolescent reform: A policy research brief*. Urbana, IL: James. R. Squire Office for Policy Research.
- National Education Association. (2010). ESEA reauthorization completes critical step in the house. Retrieved from <http://www.nea.org/home/NoChildLeftBehindAct.html>
- National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards*. Common Core State Standards Initiative. Retrieved from <http://www.corestandards.org/the-standards>
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.
- National Reading Panel. (2000). National Reading Panel releases report on research-based approaches to reading instruction: Expert panel offers its groundbreaking findings to U.S. Congress and the nation [Press release]. Retrieved from http://nationalreadingpanel.org/Press/press_rel_4_13_00.htm
- Nebelsick-Gullet, L. (2003). STAR Reading version 2.2. In B. Plake, & J.C. Impara (Eds.), *The fourteenth mental measures yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425

(2002).

Obama, B. (2011/2012). The president's record on education: Reforming No Child Left

Behind. Retrieved from [http://www.barackobama.com/record/education?](http://www.barackobama.com/record/education?source=primary-nav)

[source=primary-nav](http://www.barackobama.com/record/education?source=primary-nav)

Parris, S., & Block, C. (2007). The expertise of adolescent literacy teachers. *Journal of*

Adolescent and Adult Literacy, 50(7), 582-298.

Phelps, S. (2005). *Ten years of research on adolescent literacy, 1994-2004: A review.*

Retrieved from ERIC database. (ED 489531)

Phillips, M. (2005). *Creating a culture of literacy: A guide to middle and high school*

principals. Reston, VA: National Association of Secondary School Principals.

Pogrow, S. (2000). The unsubstantiated success of SFA. *Phi Delta Kappan*, 81(8), 596-

600.

Pogrow, S. (2001). Success for All does not produce success for students. *Phi Delta*

Kappan, 82(1), 67-80.

Pogrow, S. (2002). Success for All is a failure. *Phi Delta Kappan*, 83(6), 463-468.

Qu, Y., & Becker, B. (2003). *Does traditional teacher certification imply quality? A*

meta-analysis. Paper presented at the annual meeting of the American

Educational Research Association, Chicago, IL.

Renaissance Learning. (2009). *STAR reading resources: Pretest instructions.* Retrieved

from <http://renplace.indep.k12.mo.us/RenaissanceServer/SR/Resources/>

[Resources.aspx](http://renplace.indep.k12.mo.us/RenaissanceServer/SR/Resources/)

- Renaissance Learning. (2010a). *STAR reading computer adaptive reading test: Technical manual*. Madison, WI: Renaissance Learning, Inc.
- Renaissance Learning. (2010b). *STAR reading computer adaptive reading test: Understanding reliability and validity*. Madison, WI: Renaissance Learning, Inc.
- Rissman, L., Miller, D., & Torgesen, J. (2009). *Adolescent literacy walk-through for principals: A guide for instructional leaders*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Robinson, C. (2002). *Florida Center for Reading Research: Success for All*. Retrieved from http://www.fcrr.org/fcrrreports/pdf/success_for_all_report.pdf
- Ross, S. M., Nunnery, J. A., Goldfeder, E., McDonald, A., Rachor, R., Hornbeck, M., & Fleischman, S. (2004). Using school reform models to improve reading achievement: A longitudinal study of Direct Instruction and Success for All in an urban district. *Journal of Education for Students Placed At-Risk*, 9(4), 357-388.
- Rosen, L. (2011). Teaching the iGeneration. *Educational Leadership*, 68(5), 10-15.
- Salkind, N. (2008). *Statistics for people who (think they) hate statistics*. Thousand Oaks, CA: Sage.
- Shanahan, T. (1999). *The National Reading Panel: Using research to create more literate students*. Retrieved from <http://www.readingonline.org/critical/shanahan/panel.html>
- Slavin, R.E. (2002). Mounting evidence supports the achievement effects of Success for All. *Phi Delta Kappan*, 83(6), 469-480.
- Slavin, R. E., & Madden, N. A. (2001). Research on achievement outcomes of Success for All. *Phi Delta Kappan*, 82(1), 38-66.

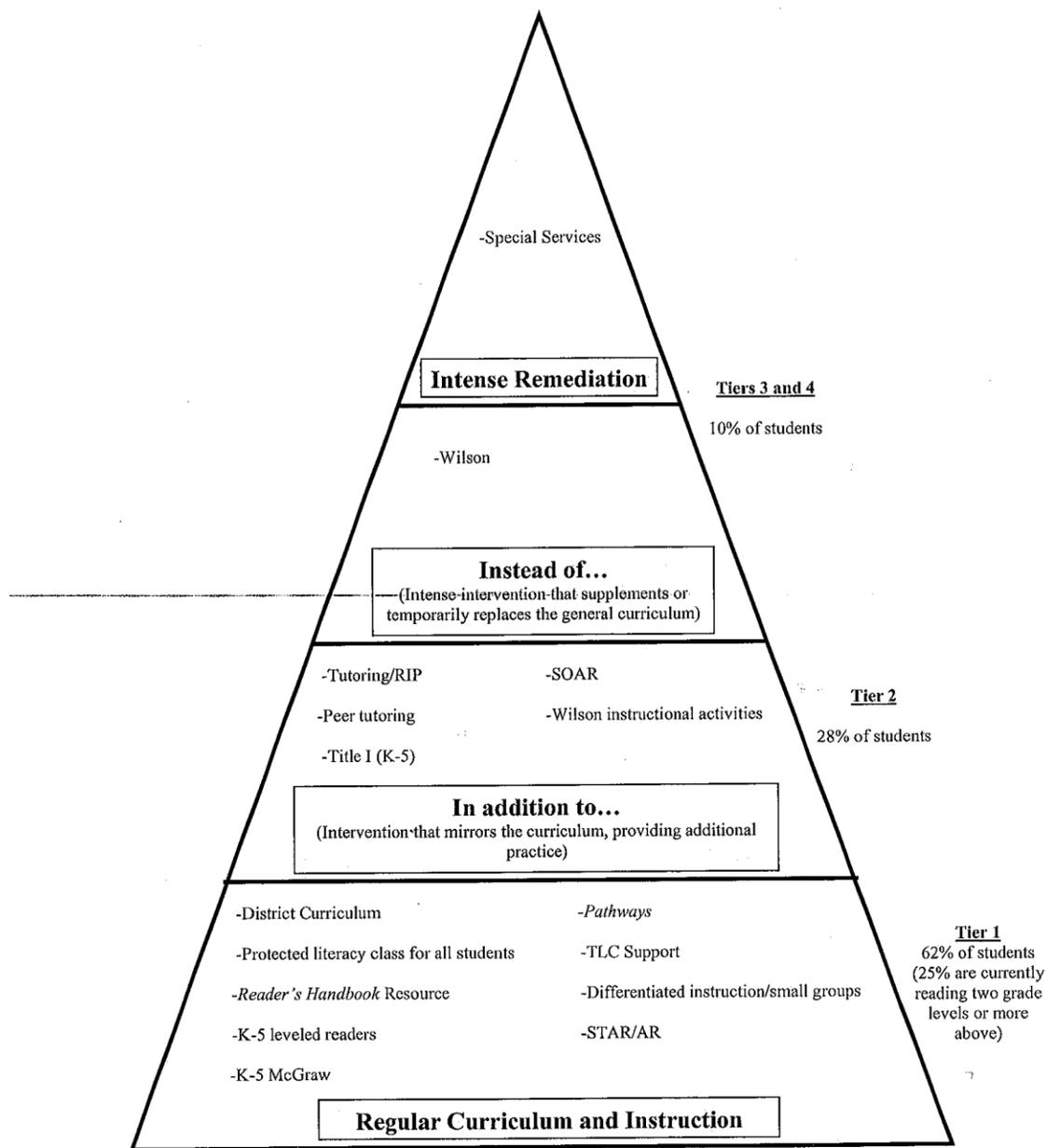
- Slavin, R., Daniels, C., & Madden, N. (2005). Success for All middle schools add content to middle grades reform. *Middle School Journal*, 36(5), 4-8.
- Slavin, R., Chamberlain, A., & Daniels, C. (2007). Preventing reading failure. *Educational Leadership*, 65(2), 22-27.
- Slavin, R., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly*, 43(3), 290-322.
- Slavin, R., Madden, N.A., Chambers, B., & Haxby, B. (2009). *2 Million children success for all*. Thousand Oakes, CA: Corwin Press.
- Smith, L., Ross, S., & Nunnery, J. (1997). *Increasing the chances of Success for All: The relationship between program implementation quality and student achievement at eight inner-city schools*. Paper presented at the 1997 annual meeting of the American Educational Research Association, Chicago, IL.
- Sparks, K. (2004). *The effect of teacher certification on student achievement* (Doctoral dissertation). Retrieved from ProQuest Dissertation and Theses database. (UMI No. 3172168)
- Strauss, S. (2002). Challenging the NICHD reading research agenda. *Phi Delta Kappan*, 86(6), 438-442.
- Success for All Foundation. (2004). *The Reading Edge: Overview*. Baltimore, MD: Success for All Foundation.
- Success for All Foundation. (2006). *Success for All: Leadership guide*. Baltimore, MD: Success for All Foundation.

- Success for All Foundation. (2011a). *Success for All Foundation: About us*. Retrieved from <http://www.successforall.org/About-Us/>
- Success for All Foundation. (2011b). *Success for All Foundation: Frequently asked questions*. Retrieved from <http://www.successforall.org/About-Us/FAQs/>
- Tovani, C. (2000). *I read it, but I don't get it*. Portland, OR: Stenhouse Publishers.
- Tovani, C. (2004). *Do I really have to teach reading?* Portland, OR: Stenhouse Publishers.
- U. S. Census Bureau. (2011). State & county quickfacts: Jackson County, MO. Retrieved from <http://quickfacts.census.gov/qfd/states/29/29095.html>
- U. S. Department of Education, Office of Elementary and Secondary Education. (2002). No Child Left Behind: A desktop reference. Retrieved from: www.nochildleftbehind.gov
- U.S. Department of Education. (2006). \$30 million in grants awarded to help struggling readers: Spellings announces first round of recipients of Striving Readers funds. *The Achiever*, 5(4), 1-6.
- US Department of Education. (2009). Federal Education Policy and the States 1945-2000: A Brief Synopsis. Retrieved from http://www.archives.nysed.gov/edpolicy/research/res_essay_intro_fedrole_states.shtml
- U.S. Department of Education. (2011). Title I — Improving the academic achievement of the disadvantaged. Retrieved from <http://www2.ed.gov/policy/elsec/leg/esea02/pg1.html>
- Vacca R., & Vacca J. (2002). *Content area reading literacy and learning across the curriculum*. Boston, MA: Allyn and Bacon.

- Veldkamp, B. (2010). Bayesian Item Selection in constrained adaptive testing using shadow tests. *Psicologica*, 31(1), 149-169.
- Vygotsky, L. (1978). *Mind in society: Development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wells, L. R., Blendinger, J., & Greene, D. (2000). An investigation of the Success for All reading program. Paper presented at the Annual Meeting of the Mid-South Educational Research Association Bowling Green, Kentucky.
- Weiler, J. (1998). *Success for All: A summary of evaluations*. Retrieved from ERIC database. (ED 425250)
- Wilson, G., Martens, P., Arya, P., & Altwerger, B. (2004). Readers, instruction, and the NRP. *Phi Delta Kappan*, 86(3), 242-246.
- Yatvin, J. (2002). Babes in the woods: The wanderings of the National Reading Panel. *Phi Delta Kappan*, 83(5), 364-369.

Appendices

Appendix A: Independence School District At-Risk Committee Documents



Note: Independence School District Response to Intervention Model, 2007.

Middle School Reading Committee
Timeline, 2007

Action	Dates/ Person(s) Responsible
Review of STAR data in order to analyze potential literacy resource needs	October, 2006 Lori, Carol, At-risk committee
Site visit to Raytown in order to view Read 180 classrooms	November, 2006 Building Representatives/Admin.
On-going discussions (What skills are our students lacking? What do our kids need to be successful in reading? Where do we go from here?)	November - December, 2006 At-risk committee, Building principals, Dr. Hartnett, Lori
Principal meeting to discuss necessity and feasibility of potential literacy resources	January, 2007
Coordination of initial reading remediation sub-committee	January, 2007
Utilize RTI model in order to identify what literacy programs/resources are currently in use or are available in the three middle schools	January, 2007 Sub-committee representatives
Based on RTI model, identify strengths and weaknesses of current literacy program (What is working? What is still needed?)	January, 2007 Sub-committee representatives
Research - Identify what local, successful (based on current MAP data) school districts are currently utilizing for resources, Call all major textbook company representatives to express needs and analyze what resources they may have that meet our district needs, Research additional programs and resources that are in existence for middle level literacy	January, 2007 Dr. Hartnett, Building principals, Lori
Add additional stakeholders to reading committee - representatives from each middle school	January, 2007
Identify those characteristics which the committee feels are required for a program to be considered "good quality" and discuss those programs and resources which fit criteria	February, 2007 Reading Committee
Connect STAR data for upcoming middle school students (2007-2008) to RTI model - What percentage of our students are going to need various interventions?	February, 2007 Reading Committee
Meet with representatives from various resources (SFA, Read 180, AMP) and view resources sent by textbook companies	March, 2007 Reading Committee - Release Day
Site visit to Rogers Middle School(KCMo) in order to observe SFA classrooms	April, 2007

Aspects of a quality literacy resource as defined by reading committee....

- Must include the five essential components as defined by DESE (phonemic awareness, phonics, vocabulary, fluency and comprehension)
- Must be research-based
- Quality assessment/diagnosis with accompanying strategies
- Prescriptive w/teacher flexibility
- On-going professional development that is specific to our needs
- Wide variety of high interest materials
- Relevant student activities
- Leveled (Remedial resources as well as those resources for students reading well above grade level)

Note: Notes distributed after a committee discussion about characteristics it required to consider a program one of “good quality” and an option for purchase.

Appendix B: IRB Application

SCHOOL OF EDUCATION
GRADUATE DEPARTMENT



Date: 7/12/2011

IRB PROTOCOL NUMBER _____
(IRB USE ONLY)

IRB REQUEST
Proposal for Research
Submitted to the Baker University Institutional Review Board

I. Research Investigator(s) (Students must list faculty sponsor first)

Department(s) School of Education Graduate Department

Name	Signature	
1. Dr. Patricia Bandré		Major Advisor
2. Margaret Waterman		Research Analyst
3.		University Committee Member
4.		External Committee Member

Principal Investigator: Pamela Lingelbach

Phone: 816-

Email: Pamela_lingelbach@indep.k12.mo.us

Mailing address: 2607 Seminole Ct., Independence, MO 64057

Faculty sponsor: Dr. Patricia Bandré
Phone: 913-344-1233
Email: patricia.bandre@bakeru.edu
Expected Category of Review: Exempt Expedited Full

II: Protocol Title

An Analysis of Student Reading Achievement During Implementation of The Success for All Foundation's Reading Edge Program in a Midwestern School District

Summary

The following summary must accompany the proposal. Be specific about exactly what participants will experience, and about the protections that have been included to safeguard participants from harm. Careful attention to the following may help facilitate the review process:

In a sentence or two, please describe the background and purpose of the research.

The purpose of this study is to evaluate the impact of the implementation of the Success for All Foundation's Reading Edge reading instructional program on student reading achievement scores for students in grades 6-8 in one school district.

A second purpose of this study is to explore the relationship between measured student reading growth and instructors' current content teaching areas to determine whether students of teachers instructing in a particular content area demonstrate greater reading growth than others.

Briefly describe each condition or manipulation to be included within the study.

There are no conditions or manipulation included in the study.

What measures or observations will be taken in the study? If any questionnaire or other instruments are used, provide a brief description and attach a copy.

Student achievement will be analyzed using Renaissance Learning Company's Standardized Test for the Assessment of Reading (STAR) Instructional Reading Level (IRL) scores. These archived scores are available through the school district database. Achievement growth in reading classes taught by teachers whose employment focuses primarily on the content areas of science, social studies, math, and communication arts will also be analyzed using STAR IRL scores. Enrollment records and teacher instructional duty assignments are available through the PowerSchool Student Information Systems district database.

Will the subjects encounter the risk of psychological, social, physical, or legal risk? If so, please describe the nature of the risk and any measures designed to mitigate that risk.

The subjects will not encounter risk of psychological, social, physical, or legal risk in this study.

Will any stress to subjects be involved? If so, please describe.

Subjects involved in this study will not experience any stress.

Will the subjects be deceived or misled in any way? If so, include an outline or script of the debriefing.

Subjects will not be deceived or misled in this study.

Will there be a request for information that subjects might consider to be personal or sensitive? If so, please include a description.

The data in this study were examined at the building and the district level. No data will be personal or sensitive in nature. Data will remain anonymous and will not be connected to individual students.

Will the subjects be presented with materials that might be considered to be offensive, threatening, or degrading? If so, please describe.

Subjects will not be presented with materials that might be considered to be offensive, threatening, or degrading.

Approximately how much time will be demanded of each subject?

No time will be demanded of each subject. The researcher is using archived score data available through district licensed Renaissance Place software.

Who will be the subjects in this study? How will they be solicited or contacted? Provide an outline or script of the information which will be provided to subjects prior to their volunteering to participate. Include a copy of any written solicitation as well as an outline of any oral solicitation.

The subjects in this study were middle school students, grades 6-8, in three buildings in the district between 2007-2010. No additional student participation will be necessary for this study. Student achievement from 2008-2010 will be analyzed according to reading teachers' primary instructional areas (math, science, social studies, communication arts). No additional teacher participation will be necessary for this study, as all enrollment and instructional assignment data will be available through archived records in the district student information system PowerSchool software.

What steps will be taken to ensure that each subject's participation is voluntary? What if any inducements will be offered to the subjects for their participation?

No additional participation will be necessary for this study.

How will you ensure that the subjects give their consent prior to participating? Will a written consent form be used? If so, include the form. If not, explain why not.

Subjects will not be contacted for this study. All scores are available to administrative users of the district-licensed Renaissance Place software.

Will any aspect of the data be made a part of any permanent record that can be identified with the subject? If so, please explain the necessity.

No data will be made a part of any permanent record as a result of this study.

Will the fact that a subject did or did not participate in a specific experiment or study be made part of any permanent record available to a supervisor, teacher or employer? If so, explain.

Archived data will be used for this study. No data will be made a part of any permanent record as a result of this study.

What steps will be taken to ensure the confidentiality of the data?

All data given to the researcher will remain confidential and will be reviewed only by the researcher and the researcher's committee. Student names will be replaced by numbers. Teacher

names will be replaced by a department title (example: Math). All records will remain in the possession of the researcher and will be stored in a locked file cabinet. After the study is complete, all paper documents will be shredded.

If there are any risks involved in the study, are there any offsetting benefits that might accrue to either the subjects or society?

There are no risks involved in this research study. The district is currently conducting ongoing program evaluation for all reading programs and interventions used K-12, so it benefitted through the triangulation of data this outside analysis and evaluation provided.

Will any data from files or archival data be used? If so, please describe.

Yes, all data used in the research study is archival and is available to district employees with password access and authorization status as administrative users of the district-licensed Renaissance Place software, a system that stores Standardized Test for the Assessment of Reading (STAR) Instructional Reading Level (IRL) scores. Additionally, archival data is also available to district employees with password access and authorization status as administrative users of the PowerSchool Student Information Software system, the program that archives master schedule information, including teacher course responsibilities and student enrollment records.

Appendix C: IRB Approval

July 13, 2011

SCHOOL OF EDUCATION
Baldwin City, Overland Park, Wichita, Topeka,
Kansas City and Lee's Summit

Pamela Lingelbach



Dear Ms. Lingelbach:

The Baker University IRB has reviewed your research project application (M-0114-0712-0713-G) and approved this project under Exempt Review. As described, the project complies with all the requirements and policies established by the University for protection of human subjects in research. Unless renewed, approval lapses one year after approval date.

The Baker University IRB requires that your consent form must include the date of approval and expiration date (one year from today). Please be aware of the following:

1. At designated intervals (usually annually) until the project is completed, a Project Status Report must be returned to the IRB.
2. Any significant change in the research protocol as described should be reviewed by this Committee prior to altering the project.
3. Notify the OIR about any new investigators not named in original application.
4. Any injury to a subject because of the research procedure must be reported to the IRB Chair or representative immediately.
5. When signed consent documents are required, the primary investigator must retain the signed consent documents for at least three years past completion of the research activity. If you use a signed consent form, provide a copy of the consent form to subjects at the time of consent.
6. If this is a funded project, keep a copy of this approval letter with your proposal/grant file.

Please inform Office of Institutional Research (OIR) or myself when this project is terminated. As noted above, you must also provide OIR with an annual status report and receive approval for maintaining your status. If your project receives funding which requests an annual update approval, you must request this from the IRB one month prior to the annual update. Thanks for your cooperation. If you have any questions, please contact me.

Sincerely,

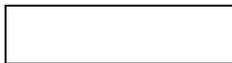
Carolyn Doelittle, EdD
Chair, Baker University IRB

UNDERGRADUATE CAMPUS | P.O. Box 65, Baldwin City, Kansas 66006
785.594.6451 | fax 785.594.2522 | www.bakerU.edu

GRADUATE CAMPUS | 8001 College Boulevard, Suite 100, Overland Park, Kansas 66210
913-491-4432 | fax 913-696-1997 | www.bakerU.edu

**Appendix D: Request to Access and Analyze Independence School District
Renaissance Place Archived Data**

Pamela Lingelbach



February 22, 2012

Dear Dr. Savidge:

I am a graduate student at Baker University and am currently pursuing a degree in District Level Leadership. I would like to request your permission to access and use archived data from the Independence School District's student STAR testing records from 2007-2010 for my dissertation. The purpose of my study is to evaluate the effect of implementing the Success for All Foundation's Reading Edge curriculum package on regular student achievement as measured the STAR.

The ISD will benefit from allowing me to analyze this data because it will provide program evaluation for a commercial curriculum that is still in use in the district. I will happily share my findings with your office and make recommendations about program use based on both the data analysis and my research of the field.

I have attached my Baker University Institutional Review Board request and approval forms to this document. For more information about the study, please contact me at the address shown above or via phone at 816-

Thank-you for your time and attention in this request.

Sincerely,

A handwritten signature in cursive script that reads "Pamela Lingelbach".

Pamela Lingelbach, NBCT

**Appendix E: Independence School District Approval to Access and Analyze
Renaissance Place Archived Data**



Dr. Elizabeth A. Savidge
Assistant Superintendent
Curriculum/Instruction/
Professional Development
Independence School District

April 7th, 2012

To Whom It May Concern:

As the Assistant Superintendent for the Independence School District, please accept this written communication as approval for Pamela Lingelbach to access the Independence School District's Renaissance Place STAR data for 2007-2010. Please feel free to contact me if you have any questions.

Respectfully,



Dr. Elizabeth A. Savidge
Assistant Superintendent
Curriculum, Instruction and Professional Development
Independence School District
Independence, Missouri
816-521-5300

Appendix F: Tukey's HSD, Year 1 Fall to Year 1 Midpoint Timeframe

Tukey's HSD, Year 1 Fall to Year 1 Midpoint Timeframe

Content		CA		MATH		SCI		SS		
		Proficiency	At / Above	Below						
		Mean	.139	.473	-.271	.697	-.107	.802	-.156	.693
CA	At / Above	.139								
	Below	.473								
MATH	At / Above	-.271	-.409							
	Below	.697		.223						
SCI	At / Above	-.107	-.245							
	Below	.802		.328						
SS	At / Above	-.156	-.295							
	Below	.693		.220						

Sign >.66

Note: CA = Communication Arts, SCI= Science, SS = Social Studies,

**Appendix G: Tukey's HSD Year 1 End of Year to Year 2 Fall
Timeframe**

Tukey's HSD Year 1 End of Year to Year 2 Fall Timeframe

Content	CA		MATH		SCI		SS		
	Proficiency	At / Above	Below						
	Mean	-.147	.473	-.735	.697	-.332	.802	-.453	.693
CA	At / Above	-.147							
	Below	.473							
MATH	At / Above	-.735	-.589						
	Below	.697	.223						
SCI	At / Above	-.332	-.185						
	Below	.802	.328						
SS	At / Above	-.453	-.307						
	Below	.693	.220						

Sign >.55

Note: CA = Communication Arts, SCI = Science, SS = Social Studies

Appendix H: Tukey's HSD, Year 2 Midpoint to Year 2 Spring Timeframe

Tukey's HSD, Year 2 Spring to Year 2 End Timeframe

Content		CA		MATH		SCI		SS	
Proficiency		At / Above	Below						
Mean		-.194	.546	-.492	.404	.006	.613	.006	.550
CA	At / Above	-.194							
	Below	.546							
MATH	At / Above	-.492	-.299						
	Below	.404	-.142						
SCI	At / Above	.006	.199						
	Below	.613	.067						
SS	At / Above	.006	-.259						
	Below	.550	.004						

Sign >.50

Note: CA = Communication Arts, SCI= Science, SS = Social Studies,

Appendix I: Tukey's HSD, Year 2 Spring to Year 2 End Timeframe

Tukey's HSD, Year 2 Spring to Year 2 End Timeframe

Content		CA		MATH		SCI		SS	
Proficiency		At / Above	Below						
Mean		-.349	.668	-.162	.586	-.125	.455	-.219	1.07
CA	At / Above	-.349							
	Below	.668							
MATH	At / Above	-.162	.187						
	Below	.586		-.083					
SCI	At / Above	-.125	.224						
	Below	.455		-.213					
SS	At / Above	-.219	.130						
	Below	1.07		.404					

Sign >.51

Note: CA = Communication Arts, SCI= Science, SS = Social Studies,