ACCELERATED READER AND ITS EFFECT ON FIFTH-GRADE STUDENTS' READING COMPREHENSION

by

Jan Shelton Nichols

Liberty University

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University

March 28, 2013

ACCELERATED READER AND ITS EFFECT ON FIFTH-GRADE STUDENTS' READING COMPREHENSION

by Jan Shelton Nichols

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University, Lynchburg, VA

March 28, 2013

APPROVED BY:

LISA REASON, PhD, Committee Chair

ANDREA RAY, PhD, Committee Member

JESSICA MCCLUNG, EdD, Committee Member

Scott B. Watson, PhD, Associate Dean, Advanced Programs

ABSTRACT

Jan Shelton Nichols. ACCELERATED READER AND ITS EFFECT ON FIFTH-GRADE STUDENTS' READING COMPREHENSION. School of Education, March, 2013.

Schools in the United States have been using the Accelerated Reader (AR) program since the mid-1980s. A search of the literature related to the effectiveness of the AR program revealed that many of the studies were conducted more than a decade ago, and a large number of those studies failed to utilize a control group to provide comparative data. Researchers and educators have examined the use of the AR program, whose parent company is Renaissance Learning, for more than 20 years, yet there is still little definitive data on whether the system positively affects student reading comprehension and motivation to read. This causal-comparative design, in the form of an ex post facto study, examined two groups of fifth-grade students to determine if the addition of AR showed statistically significant effects on their reading comprehension. Data were gathered from both an experimental and a control group; the control group utilized the school system's literacy plan for reading instruction, while the experimental group utilized the literacy plan along with mandatory use of AR. Independent t-tests were used to determine if the treatment (AR) had any effect on reading comprehension by group or by gender. The results from the independent t-tests showed no statistically significant effect for reading method by group and no statistically significant effect for gender by group.

Acknowledgments

This dissertation is the final product of a long journey, and the joys and pains that go along with a process such as this cannot be managed alone. My spiritual life has been forever changed due to the constant prayers for strength and wisdom and the resulting continuous flow of blessings from my Lord and Savior. I have also been carried by my husband's love through many long and stressful days. Al's unceasing faith in my ability to persevere and his altering of his life to accommodate my school requirements are perfect examples of why I consider him a gift directly from God's hand. Our children, Jonathan (and his wife Carla) and Amberly (and her husband Steve) have been amazing during this time as they stepped back and waited for their mother to emerge from her office and finally spend time with them. We began this doctoral journey with no grandchildren and have finished it with three—Cristian, Isabella, and Ava. What an amazing legacy and testament of the love between two people who married young and gave their children back to God for His purposes.

I am so very thankful to those at Liberty University who walked me through this dissertation process and guided me to a finished product. As my dissertation committee chair, Dr. Lisa Reason provided excellent guidance and support through calm and reassuring feedback. Dr. Andrea Ray also guided and encouraged me with focused feedback and encouragement. Dr. Jessica McClung, my local committee member, provided a shoulder to cry on and a steady stream of much needed cheerleading—what an answer to prayer! Dr. Scott Watson provided amazing guidance through the statistical gauntlet of this journey, and his patience and advice have been greatly appreciated. And… I cannot forget my special coworker, Randy Queen, who functioned as my daily

advisor and sounding-board from the beginning of the doctoral process (it is his turn next!). He also took on the role of "research assistant" to help with the final phase of the project. Randy is proof that God puts special people in your path for a purpose. Randy's counseling skills were used on many occasions!

This completed dissertation is the achievement of a lifelong professional dream, but the true benefits seen from this adventure are those that have drawn me closer to my Lord and Savior and to my husband. The rest of my days will be spent thanking them both for their ever-present, supportive presence in my life. God is good.

Table of Contents

List of Tables

CHAPTER ONE: INTRODUCTION

Introduction to the Problem

The purpose of this study was to add to the body of literature that attempts to analyze the effectiveness of Accelerated Reader (AR) on reading comprehension. A search for literature related to the effectiveness of AR revealed two resources: the U.S. Department of Education's (USDE) What Works Clearinghouse (WWC) and the studies listed as evidence of the effectiveness of AR on the Renaissance Learning (2012) website. The WWC's role in education is to review studies related to best practices in classrooms and identify those studies whose processes meet their stringent criteria for effectiveness. After close review of both sites, it was determined that this research study would examine those studies that were included on the WWC (2010) website as well as the Renaissance Learning (2012) site.

Upon review, it was found that there were two significant gaps in the literature available for educators' use in determining the effectiveness of the AR program. First, the bulk of the studies were completed more than 7 years ago. Second, most of the studies listed on the Renaissance Learning website did not involve scientifically managed conditions. A review of all 166 studies revealed that the majority of the titles on the Renaissance Learning list fit into one of two categories. Those that were listed as studies were either anecdotal articles reporting the change in student performance after simply adding AR to the school program without benefit of a comparison group or studies that used a specific group with a pre- and a posttest but no comparison group. This study involved a specific group of fifth-grade students who consistently participated in the AR

program along with the school system's literacy plan compared to a second group of fifth-grade students who participated in only the literacy plan for reading instruction.

Teachers are constantly seeking a strategy that will increase student motivation to read as well as increase their comprehension. The demands placed on teachers in elementary classrooms in the United States are continuing to increase due to pressure from the federal government through legislation such as the No Child Left Behind Act of 2001 (NCLB; Allington, 2005; Blankstein, 2010; Brozo, 2010; Hansen, Collins, & Warschauer, 2009; Hilden & Pressley, 2007). This legislation requires that schools work diligently to raise the reading scores of all students to grade-level expectations by 2014 (NCLB, 2002). As children progress through elementary school, there are fewer instructional minutes available to spend on the task of learning to read (Lane, Little, Redding-Rhodes, Phillips, & Welsh, 2007). In Virginia, students begin their Standards of Learning (SOL) tests in third grade with a total of four high-stakes tests: reading, math, social studies, and science. Teachers are constrained by the burden of teaching the content for these tests; consequently, little time is available for continued basic reading instruction.

It is this lack of time for reading instruction, yet the ever-increasing accountability for teaching students to be good readers, that causes schools to search for quick and easy strategies to boost their students' comprehension scores (Meier, 2004). Thousands of schools have turned to programs such as AR (Hansen et al., 2009; D. W. Johnson & Johnson, 1999; Nunnery & Ross, 2007) in search of that instant, simple solution to their students' complex reading difficulties (Blankstein, 2011).

The parent company of Accelerated Reader, Renaissance Learning, advertises through its website that the AR program will increase students' reading comprehension levels. According to its website, Renaissance Learning (2012) asserts that AR is "the most cost-effective and successful reading software of all time." Although Renaissance Learning states that the AR program is fully supported by the USDE through the WWC, it is noted that of the 318 studies reviewed by the WWC, only one study—Bullock (2005)—was considered to meet its evidence standards without reservation. A study conducted by Nunnery and Ross (2007) was found to meet those standards, with reservation. The other 316 studies were unable to be used for evaluation due to failure to meet protocol or other WWC standards (WWC, 2010). Consequently, the WWC concluded that AR did not show statistically significant effects for fluency, nor did it show any statistical significance for comprehension (which is a key selling point for AR).

Background of the Study

AR was initially a product of Advantage Learning Systems, whose name changed to Renaissance Learning in 2001 (Investors' Business Daily, 2011). The program was developed in 1984 by Terry D. Paul and was designed to motivate reluctant readers to increase their independent reading time as well as their comprehension of what they had read. The program has become one of the most widespread computer-based reading programs ever released and is found in thousands of schools around the world (Hansen et al., 2009; D. W. Johnson & Johnson, 1999; Nunnery & Ross, 2007).

Using the AR program requires that students first take a computerized test to determine their AR reading level. This test, the Standardized Test for Assessment of Reading (STAR) test, is also a Renaissance Learning product. The multiple-choice test

takes approximately 20 minutes to complete. The system begins the test setting with practice questions and, if the student is successful, provides questions that will be scored based on the student's perceived ability level (Renaissance Learning, 2010). Once the test is completed, the system will provide a variety of information to the teacher, including an AR reading range. This range is referred to as the student's zone of proximal development (ZPD).

Once the ZPD is determined, the teacher meets with the student and provides a list of books available within the student's range. The student is then able to read books on the AR list that are available in his or her school library. After reading a book, the student can take a 5- to 20-question multiple-choice test that is designed to determine the student's comprehension of the book. Students are considered performing at their designated ZPD level when their scores are consistently at or above 90–100% (Renaissance Learning, 2007). Although Renaissance Learning states that 40 tests are available for teachers to review to determine student reading progress, it is this researcher's experience that the only report used is the student report that indicates which tests have been taken and the score on those tests. After spending 20 years in six different schools in two states (all of which used AR), this researcher was not aware of these additional resources until conducting this study.

Problem Statement

The impetus for this study was the continual observation of teachers who depend on the AR program as a scientifically proven strategy to improve student motivation to read as well as their comprehension when reading. There is a definite need for studies that examine AR use in a variety of settings to determine if the program delivers on the

company's promises (Hansen et al., 2009; R. A. Johnson & Howard, 2003; Nunnery, Ross, & McDonald, 2006; Schacter, 2000). On its website, Renaissance Learning (2012) states that its program has been added to over 70,000 K–12 schools, which indicates that 70,000 libraries have invested in labeling books with AR levels and have purchased tests for each of those labeled books. That is a tremendous investment to make without benefit of true, unbiased research to support or refute AR's claims. This study examined a specific setting to add to the literature base to assist in the determination of the effect AR may or may not have on student reading progress.

The review of literature for this study revealed that other researchers have expressed doubts about the use of AR as a major part of a classroom's reading program and the need for more current studies (Hansen et al., 2009). The claims made by Renaissance Learning (2012) did not materialize in many of the studies, and some studies indicated that students who participated in the program did not progress as predicted by the program's makers (Bullock, 2005; Cuddeback & Ceprano, 2002; Esposito & Smith, 2006; Facemire, 2000; Florida Center for Reading Research [FCRR], 2006; Jenkins & Terjeson, 2011; Macaruso & Walker, 2008; Nunnery & Ross, 2007; Nunnery et al., 2006; Oppenheimer, 2007; Pavonetti, Brimmer, & Cipielewski, 2003; G. Thompson, Madhuri, & Taylor, 2008). It was surprising that several of these studies that refuted the benefit of AR were listed on the Renaissance Learning site as supportive of its program.

AR also noted on its website that it has the support of the WWC, a USDE program that reviews best practices that are research-based. But according to the WWC's 2010 report, there is a distinct lack of studies that indicate that the use of AR has a positive impact on student reading performance. In fact, the WWC found that only one

study—Bullock (2005)—out of 318, was acceptable as a study that used a pre- and posttest as well as a comparison group under quasiexperimental conditions. An unexpected finding in the study was that it did not support AR as a viable tool for increasing student reading performance. Bullock found that AR's lack of an instructional component that is explicit in nature limits the program's effectiveness. Bullock believed as this researcher does: There is a great need for more quasiexperimental studies that seek to identify the depth of the effectiveness of the AR program.

Bullock's study was conducted in 2005. The WWC's last report of effectiveness was completed in 2010, and most of the studies reviewed in its report date from 1999 to 2005. The lack of a large pool of current studies from which to draw data limits the ability to generalize the data to the possible effectiveness of the program for students in classrooms today. An additional concern related to the use of data that may be outdated is the fact that many of the studies were not conducted under circumstances that would allow the elimination of competing interventions in the classroom that may have affected the outcomes (WWC, 2010).

Purpose of the Study

The purpose of this quantitative study was to provide current data gained under quasiexperimental conditions. This causal-comparative study was designed to examine the possible effects of the use of AR when one group of fifth-grade students used the school system's literacy plan along with mandated use of the AR program throughout the school year. The comparison group of fifth-grade students used the school system's literacy plan without the AR program. To determine if the addition of the AR program

affects reading comprehension, 2011 SOL reading scores for fourth grade were used as the pretest and the 2012 SOL reading scores for fifth grade were used as the posttest.

Research Questions and Hypotheses

Research Question 1: What is the difference between the 2012 Virginia Standards of Learning reading scores in fifth-grade students who used the system literacy plan along with mandatory use of Accelerated Reader as compared to the scores of fifth-grade students who used only the literacy plan?

Null Hypothesis 1: There will be no statistically significant difference between the scores of fifth-grade students who used the system literacy plan along with mandatory use of Accelerated Reader when compared to the scores of fifth-grade students who used the literacy plan without mandatory use of AR as evidenced by their performance on the 2012 fifth-grade Standards of Learning reading test.

Research Question 2: What is the difference between the 2012 Virginia Standards of Learning reading scores by gender for fifth-grade students who used the system literacy plan along with mandatory use of Accelerated Reader as compared to the scores of fifth-grade students who used only the literacy plan?

Null Hypothesis 2: There will be no statistically significant difference between the scores of female fifth-grade students who used the system literacy plan along with mandatory use of AR when compared to the scores of female fifth-grade students who used the literacy plan without mandatory use of AR as evidenced by their performance on the 2012 fifth-grade Standards of Learning reading test.

Null Hypothesis 3: There will be no statistically significant difference between the scores of male fifth-grade students who used the system literacy plan along with

mandatory use of AR when compared to the scores of male fifth-grade students who used the literacy plan without mandatory use of AR as evidenced by their performance on the 2012 fifth-grade Standards of Learning reading test.

Significance of the Study

Renaissance Learning (2012) has stated that more than 70,000 schools use the AR program. This means that millions of students attend schools that encourage their use of AR each day. It is vital that administrators and teachers be provided with evidence of AR's effects in various settings in order to determine if they are serving students' best interests by requiring their participation in a program that, as the research has shown, does not always deliver what it advertises (Hansen et al., 2009; R. A. Johnson & Howard, 2003; Nunnery et al., 2006; Schacter, 2000). Reading is a fundamental skill that greatly improves a student's day-to-day existence (McCardle & Chhabra, 2004). The value of developing as a strong, solid reader can never be underestimated. On his website, Walter (2010) articulated his vision of reading by stating that the value of reading and all that it brings to a student's life is "incalculable." He compared the value of reading with the cost of ignorance that is seen on the streets each day as nonreading students are lost to a life of desperation and hopelessness. Blankstein (2010) noted that the original mission of public schools was to provide access to a basic education for all children. The current mission, though, is for all schools to achieve 90–95% success with all students.

Those educators who work with children each day understand the importance of teaching all students to read, and they labor daily to meet the needs of those who struggle (Allington, 2005; Ferguson, Currie, Paul, & Topping, 2011; Guthrie et al., 2004). But while these same teachers are inundated with strategies and interventions, they remain

confused as to which are truly the best research-based practices (Ferguson et al., 2011). And still some students fail to gain the skills necessary to navigate the texts they encounter in the classroom each year. The specter of future failure for students weighs heavily on teachers and administrators as they recognize that if students are not proficient readers at the end of the third grade, research shows they may never gain those skills (Benner, Nelson, Ralston, & Mooney, 2010; Blankstein, 2010). It is research such as this current study that will give teachers and administrators insight into AR's effectiveness in a specific setting with a specific population.

There is also a more routine reason to help students gain reading proficiency through effective practices. Under the accountability requirements of NCLB (2002), school systems are obligated, in order to continue to receive federal funds under Title I, to show that students present adequate yearly progress (AYP) at the end of each school year. AYP is determined through subgroup success on annual measurable objectives (AMOs), and this growth is to be shown through valid and reliable end-of-year testing that has been approved by the USDE, which, in Virginia's case, is the SOL tests.

The threat of sanctions and loss of funding have forced educators to seek programs that can deliver quick fixes for struggling students, with little regard to best practices and individual needs (Pitcher, Martinez, Dicembre, Fewster, & McCormick, 2010). The purpose for instruction has become not the education of the child but the passing of the test (Meier, 2004). This study adds to the resources available for educators to assist in determining the most effective use of limited instructional funds.

Definition of Terms

Adequate yearly progress (AYP): A measure of year-to-year student achievement on statewide assessments. Each state comes up with its own definition on what it means to make AYP. Definitions must answer three questions: the percentage of students that must be proficient or above when tested in reading and mathematics (yearly in grades 3-8 and once in high school); whether or not at least 95 percent of students in those grades participated in the assessments; and, the additional academic indicator (e.g., graduation rates for high schools) that will be measured. (USDE, 2012)

Annual Measurable Objectives (AMOs): A goal that a state sets each year to define a minimum percentage of students who must meet or exceed standards on its academic assessments. Each state's AMOs are applied consistently throughout the state for all public schools, districts, and subgroups of students. All students must be proficient in reading/language arts and mathematics by 2013-14. (USDE, 2012)

Grade equivalent: Refers to the scale used for AR leveling; the scores range from 0.0 to 12.9+. Student STAR test performance is compared to other students nationally to obtain this level (Renaissance Learning, 2010).

Standards of Learning (SOL) tests: Developed by Pearson Testing for the state of Virginia in order to meet the requirements of NCLB. These standardized tests are given at the end of each course or at the end of each school year to measure student achievement in English, mathematics, science, and history/social science (Virginia Department of Education, 2010).

Zone of proximal development (ZPD): Refers to the difference between what a child can do independently and what a child can do when assisted by a teacher. The

theory follows the premise that the child will take on more of the problem solving as the adult gradually removes support (Vygotsky, 1986).

CHAPTER TWO: REVIEW OF THE LITERATURE

Reading is a vitally important skill for all students. But there is a greater reason than simply academic achievement for students to leave elementary school with strong reading skills. Daniel et al. (2006) looked at the relationship between reading problems of adolescents and suicidality and dropout potential. Their chilling results revealed that students who cannot read well at the end of elementary school were more likely to think about suicide and to attempt suicide. The dropout rate was significantly higher for poor readers as well (Blankstein, 2010; Denton & Al Otaiba, 2011).

Denton and Al Otaiba (2011) also found that if a student leaves elementary school as a poor reader, he or she is destined to remain a poor reader, without significant interventions. This lack of reading skills is a highly predictive indicator that a student will become a school dropout (Benner et al., 2010). These studies advocate for a system of reading instruction quite different from that of AR. They advocate for a system that does not simply assess how well a student can condense his or her understanding of a book to the correct answering of 5–20 multiple-choice questions but an assessment system that provides teachers with a clear picture of the reading deficits of the student in order to prescribe the appropriate interventions (Benner et al., 2010; Denton & Al Otaiba, 2011; Moore, Bean, Birdyshaw, & Rycik, 1999; Pitcher et al., 2010).

This causal-comparative study was designed to provide another lens through which administrators can view the effectiveness of the AR program in a specific setting. A review of the literature related to AR revealed a dearth of studies conducted within the last 5–7 years. This shortage of current data makes it difficult for educators to determine whether AR is the most effective investment of their instructional funds. An additional

concern surfaced as the literature was reviewed on the Renaissance Learning (2012) website. Of the 166 studies reviewed on the site, only 34% could be considered studies that were conducted under quasiexperimental research conditions. The remaining 66% of the studies were actually reports from teachers and administrators or in-house reports generated by the parent company of AR, Renaissance Learning. Several researchers questioned the objectivity that could be found in studies conducted by AR creator, Terrence Paul, or studies funded by Renaissance Learning (Hansen et al., 2009; R. A. Johnson & Howard, 2003; Nunnery & Ross, 2007). Other researchers have also questioned whether the research listed by AR could stand up under closer examination (Nunnery et al., 2006; Oppenheimer, 2007; Pavonetti et al., 2003).

It is vital that the reading performance of fifth-grade students be monitored closely to ensure that they are prepared for the rigors of middle school. But teachers at the fifth-grade level are not typically trained as teachers of reading (Hilden & Pressley, 2007). Upper elementary teachers tend to be more content-specific and able to teach the features of reading such as author's purpose, genre, and vocabulary, but they are not trained to be beginning reading teachers. This lack of skills as well as a lack of time to gain those skills creates the perfect environment to make a program such as AR attractive to both teachers and administrators (Lane et al., 2007).

As mentioned previously, the motivation for all instructional decisions made in U.S. public schools can be found in the rebirth of the Elementary and Secondary Education Act (ESEA). The ESEA was reauthorized under President George W. Bush and is now known as NCLB; it directs that students must meet AMOs each year as

determined by a percentage that increases with each school year. The subgroups measured each year are as follows:

- Asian/Pacific Islander
- American Indian/Alaskan Native
- Black (not Hispanic)
- Hispanic
- White (not Hispanic)
- Multiracial/Other
- Limited English proficiency (LEP)
- Students with a disability (SWD)
- Disadvantaged (those qualifying for free or reduced lunch)

For school year 2010–2011, all subgroups were required to meet an AMO percentage of 85 in math and 86 in reading. According to the Virginia Department of Education (2012a), at this level, the percentage of schools in the Commonwealth of Virginia that met this benchmark was 38% (schools) and 3% (school divisions)—an extremely low pass rate for this state. Barring any changes by President Obama, NCLB (2002) requires that all students perform at grade level by 2014. Failure to meet this monumental goal comes with the realization that a school or system could face sanctions and loss of federal funding.

The underlying funding source for NCLB is Title I, which provides much of the remediation and staffing money used to serve disadvantaged students. It is the threat of withholding this funding that keeps America's schools grasping for interventions and strategies that will help students who are disadvantaged, minority, limited English

speakers, or identified for special education—the four subgroups whose progress is of the most concern (Meier, 2004).

To monitor the success of these specific subgroups, the federal government requires each state to develop rigorous standards for content and instruction. To measure student success, each state must also develop end-of-year or end-of-course content-specific, standardized tests for every student in Grades 3–8. These tests must be approved by the USDE in order for states to remain qualified for funding (NCLB, 2002).

Prior to NCLB, schools utilized many tools to assess student learning each year. However, the use of a one-time standardized test as the final measure of a student's mastery of a year's schooling has changed the face of classroom instruction. Teachers' emphasis has changed from content that is meaningful and enriched to increasing test scores on basic content (Kohn, 2004).

The system in which this study took place uses quarterly benchmark assessments that mimic the SOL tests. This means that elementary students will have participated in at least three high-stakes multiple-choice tests before the SOL test is given at the end of the school year. When so much emphasis is placed on rote memory of content knowledge as opposed to problem solving and assimilation of information, the ability of teachers to work with students where they are academically and take them where they need to be becomes a fading memory (Hilden & Pressley, 2007; Lane et al., 2007). Many classrooms are no longer learner-centered, they are testing-centered. However, it is the learner-centered environment that promotes true learning (Bransford, Brown, & Cocking, 2000). This forced change in focus has pushed teachers and administrators to seek strategies that work quickly and economically. The AR program has promised to assist

teachers in their efforts to not only help students become better readers but also create in students a hunger for reading in the process. But Blankstein (2010) cautioned that a common reaction of educators when faced with a seemingly insurmountable challenge is to seek a quick and easy solution—regardless of the complexity of the challenge. Alvermann (2003) also noted that those in education need to alter their mindset that there is a "magic bullet" (p. 2) that will simply fix students' reading problems and recognize that improvement in reading achievement will only occur when teachers alter their teaching methods to meet individual student needs.

Understanding the Accelerated Reader Program

Information regarding the AR program is readily available on the Renaissance Learning (2012) website. The AR program began in the mid-1980s as the brainchild of Judi and Terrance Paul. AR was created to fill a void for students by providing short quizzes that could be taken after reading a book. Students would not only receive immediate feedback on their success on the quiz but they would also begin to accumulate points on each book, which could result in additional prizes or recognition. The hope was that this feedback and the extrinsic motivators would increase students' desire to read as well as increase their comprehension of what they had read.

There is a strong following of dedicated educators who feel that AR produces a love of reading in every student who uses the program. This fervor was described by Oppenheimer (2007) in his portrayal of Renaissance Learning's large conferences as having the feel of "religious revival meetings" (p. 25). But it is important to note that, although AR may have humble beginnings with a pure pedagogical mission to increase children's love of reading, the company is now a multi-million-dollar corporation. AR is

also a very lucrative business venture. According to the investment research site Investors' Business Daily (2011), Renaissance Learning's third-quarter report showed that it had performed better than 86% of the stocks in Investors' Business Daily's database. Profits grew by 20%, and the latest growth was set at 15%, to $35.72 million—with a profit margin of 26.6 cents per dollar. It should also be noted that in August 2011, Renaissance Learning was sold to Permira (a private British equity company) for $440 million. It was the intention of this study to add to the research that may help determine if this huge increase in financial gain is due to a program that truly increases students' reading achievement or the results of a program with a successful advertising campaign.

AR's infancy coincided with the addition of desktop computers to classrooms as well as the addition of school computer labs. The AR program took advantage of this new technology and soon became the flagship product of Advantage Learning Systems—whose name changed to Renaissance Learning in 2001. AR quickly became one of the most-used computer reading programs in the country (Hansen et al., 2009; D. W. Johnson & Johnson, 1999; Nunnery & Ross, 2007).

AR costs. Adding AR to a school's instructional program requires a long-term financial commitment. Initially, schools purchased site licenses at a cost of approximately $3,000, with an additional yearly cost of $4 per student. Schools also purchased the STAR program, which provided the tests to determine a student's reading level. Initial STAR costs were approximately $1,300, with a yearly cost of approximately $250.

Once the system was loaded onto computers, librarians and classroom teachers began the task of determining which books in their existing libraries were on the AR list, labeling those books, and purchasing the available AR tests at a cost of $3 per test. When

books are labeled and tests purchased (which is an ongoing process), students are able to begin taking tests on books they have read. This researcher's school spends at least $300 per year to purchase AR tests.

AR leveling system. Accelerated Reader books are leveled according to the Advantage–TASA Open Standard (ATOS) readability formula. This system provides reading levels divided by grade level for entire books as opposed to simply leveling passages from books (Milone, 2008). Renaissance Learning worked with a number of readability experts, including Touchstone Applied Science Associates (TASA), who played a key role in the development of the new readability formula. The experts developing ATOS determined the characteristics of texts that could be objectively used as the basis for the formula: sentence length, word length, and word frequency. Other attributes were examined as well, such as the number of dependent and independent clauses and the length of the clauses, verb forms, pronoun use, paragraph length, and the ratio of unique words to the total number of words (Milone, 2008). This reading level was then utilized to determine the point value for each book using the following formula (Paul, 1996): AR Points = (10 + Reading Level) x Words in Book/100,000.

The reading level for each AR book is placed on the spine of the book for easy identification in library or classroom collections. This label also provides a less invasive manner in which teachers can monitor the levels of the books that students are reading.

Motivation and incentive debate. The original incentives for AR were in the form of immediate feedback after taking the quiz as well as the accruing of points from each test taken and passed. The incentives have moved beyond the simple idea of giving students immediate feedback. Schools that use AR provide some type of extrinsic

motivator to encourage students to meet their point goals. This researcher has observed PTAs that provide huge end-of-year celebrations with carnival-like attractions that ostracize students who did not meet their AR goal for the year. There have been less complicated and less offensive recognitions, such as bookmarks and pencils. Most schools provide pizza or popcorn parties for students who meet their goals. This emphasis on extrinsic motivation has been studied by other researchers to determine the effectiveness of essentially paying students to read. And, unless the studies examining the benefits of extrinsic motivators were conducted as a part of Renaissance Learning, the findings tend to establish a negative association between motivation and extrinsic rewards.

Dai and Wang (2007) found that student beliefs related to their reading ability greatly impacted their reading performance. Extrinsic motivators did not change this cognitive belief. Guthrie et al. (2007) as well as Logan, Medford, and Hughes (2011) found that internally motivated students had higher reading comprehension levels due to increased engagement with the text. This engagement came from the use of meaningful texts and teacher intervention in the use of the texts—not extrinsic motivators. Katz and Assor (2006) found that the strongest motivator for young readers was that of having a choice of what they wanted to read. The researchers referred to this as the "expectancy value model (EVM)" (Katz & Assor, 2006, p. 430) of achievement motivation. The EVM focuses on persistence of student effort in their reading and the social-psychological influences in place due to this level of control. The EVM fit well with Wolters's (2003) study of cognitive models of motivation that showed that individuals

purposefully act to complete an activity or task through self-regulation. This self-regulation is a cognitive process that is internally motivated.

Other researchers questioned the benefit of utilizing external motivators to increase student performance. In the 2008 study by G. Thompson et al., it was determined that the competitive environment created by the use of incentives as motivators can establish a false confidence in any increase in student reading; the dilemma created was that of determining whether students were reading because of increased desire or to earn points and prizes. In contrast, Bransford et al. (2000) found that environments that are learner-centered, where reading instruction is designed to build on student interest, strengths, and needs, tended to create an authentic motivation to read and learn. Ryan and Deci (2000) noted that intrinsic motivation resulted in higher quality learning and allowed for creativity as opposed to extrinsic motivation, which they described as a "pale and impoverished (even if powerful) form of motivation" (p. 55).

Motivation was also discussed in several studies in relation to self-determination theory (Deci & Ryan, 2008; Katz & Assor, 2006; Patall, Cooper, & Robinson, 2008; Ryan & Deci, 2000). Self-determination theory refers to the amount of control or autonomy an individual believes he or she has over an expectation. According to Deci and Ryan (2008), autonomous motivation utilizes both intrinsic and extrinsic motivation. An autonomously motivated person has developed buy-in for a specific activity, and although there may be extrinsic motivators, the person believes he or she has some control over the effects of the motivators. A controlled-motivation person believes that his or her actions are externally controlled through reward or punishment; thus, compliance is given to avoid shame or public comparison. Deci and Ryan stated that

students who are autonomously motivated are more psychologically healthy and tend to perform more effectively.

In *Getting Results With Accelerated Reader* (Renaissance Learning, 2007), hereafter referred to as the *AR Guide*, teachers are advised to refrain from using incentives to encourage reading, yet when perusing the articles in *Extraordinary Educators*, a Renaissance Learning publication, many articles describe the incentive programs used by schools throughout the nation. A particular article in the May 2011 publication discussed the use of incentives for students who score 100% on AR tests as well as school-wide incentives for high achievers in AR. However, studies such as Wang and Guthrie's (2004) showed that providing extrinsic motivation related to improving student reading performance may show positive results, but the increase is not initiated by increased interest in reading but by an attempt to gain a socially valued reward.

Extrinsic motivation is an area that has been researched extensively, yet there continues to be disagreement related to the advantages of providing the variety of extrinsic motivators that seem to go hand-in-hand with an AR program. One such study, conducted by Melton et al. (2004), examined fifth-grade achievement growth and found that when AR incentives were provided, students met their reading goals, but once the incentives were removed, so did reading as a voluntary activity. Esposito and Smith (2006) interviewed third-grade students who reported that they would not participate in AR if they did not receive rewards for their efforts. Stanfield's 2006 study also looked at third graders and found that the rewards for participation in AR actually caused a decrease in students' positive attitudes toward reading. Husman, Brem, and Duggan (2005) examined AR's effect on goal orientation and found that even with the use of

incentives, students' motivation to read remained neutral. Park (2011) studied motivational constructs and found that educators should initially seek to emphasize internal motivation. Once students have acquired internal motivation for reading, external motivators may be added. The discrepancy between what researchers have found and what Renaissance Learning advertises begs closer scrutiny through future research.

The disputable benefits of AR. The website for AR is found as a link on the Renaissance Learning (2012) website, and it boasts that AR is the "most cost-effective and successful reading software of all time." An examination of the "Overview" link finds several bold promises listed for those who utilize the AR system. The website states that AR's technology will help teachers

- Make essential reading practice more effective for every student.
- Personalize reading practice to each student's current level.
- Manage all reading activities including *read to, read with,* and *read independently.*
- Assess student's reading with four types of quizzes: Reading Practice, Vocabulary Practice, Literacy Skills, and Textbook quizzes.
- Build a lifelong love of reading and learning (Renaissance Learning, 2012).

According to the Renaissance Learning (2012) website, the aforementioned list can be supported through the massive list of studies that prove the claims as to the value of AR at every grade level in public schools. This researcher followed the link and used the search tool to request studies at all levels related to AR. The request included K–12 schools, all subgroups, and all economic levels; the search yielded 166 studies.

Upon review of the 166 studies, it was found that 65, a full 39%, were anecdotal reports submitted by teachers or administrators based on their experiences with the product in their schools. These studies included site-specific evidence related to their students' performance before AR and how they performed after AR had been implemented, without controlling for any other variables that might have influenced their achievement.

Of the remaining studies, 44 (27%) were articles and summaries produced or funded by AR's parent company, Renaissance Learning and 23 (14%) were conducted by researchers outside of Renaissance Learning, with the exception of two studies completed by Keith Topping, who has coauthored studies with AR creator Terrence Paul and wrote the foreword in a 1996 study by Paul.

The remaining 34 studies were those reviewed by the WWC, a department created by the USDE to research vendors' claims of the effectiveness of their products. Once the research is reviewed, the WWC then publishes a list of products that have met its stringent standards. AR states that its product is endorsed by the WWC for effectiveness; however, when the AR section of the WWC site is reviewed, 318 studies are listed, but only one study—Bullock (2005)— was found to meet the criteria, and another study—Nunnery and Ross (2007)—was found to meet standards, with reservation. The WWC found that, in these two studies, AR has no discernible effects related to adolescent readers' fluency or comprehension. The findings of the WWC along with this researcher's experience with AR were the motivating force behind this study to help determine if the extent of student benefit outweighs the increasing cost of the program.

While reviewing each of the studies listed on the search results, an additional area of interest arose: Several of the studies were simply lists of studies—not meta-analyses—but lists of studies that were already shown on the site individually. Several of the studies had been cited more than once, and to add to the confusion, studies that had been done by multiple researchers were listed by the original authors' names and were then found rewritten under the names of the other researchers utilizing the same data. The fact that the research listed as evidence of AR's success seemed questionable made it difficult to keep an unbiased opinion while attempting to substantiate the claims for the program.

The first step: determining reading levels through the STAR Reading assessment. The AR website provides resources for teachers to assist them in setting up their classrooms for AR use. According to the *AR Guide* (Renaissance Learning, 2007), AR is a process by which a student can test his or her knowledge of a recently read book through a brief, multiple-choice quiz. Before books are chosen, each student takes an assessment through the companion system, Standardized Test for Assessment of Reading. STAR is a norm-referenced as well as a criterion-referenced assessment that aligns suitably with AR. Grade K–2 students respond to 25 vocabulary-in-context questions, while Grade 3 and older students respond to 20 vocabulary-in-context questions with the addition of five text passages. This assessment takes 10–15 minutes and increases in difficulty as the student correctly answers questions (Renaissance Learning, 2010).

When the STAR test is completed, the teacher is able to access a reading range for each student. This range is called a zone of proximal development and utilizes grade-level distinctions for book levels. This ZPD is loosely based on Vygotsky's work related to a student's interaction with a teacher. In Vygotsky's ZPD, the teacher retreats as the

student gains mastery (Knapp, 2008; Kozulin, 2002; Kravtsova, 2009; Niewolny & Wilson, 2009; Vygotsky, 1978). In the case of AR's use of the term, the teacher uses the range provided by the STAR assessment to guide the student in choosing books at his or her reading level with the hope that the student will eventually acquire appropriate book selection skills.

To begin the AR reading and testing process, the *AR Guide* (Renaissance Learning, 2007) instructs teachers to use the students' ZPD to make book recommendations. The ZPD is gained from the STAR test that is taken periodically throughout the year. Teachers are encouraged to set point goals wherein the student either works toward a specific point goal or the points are recorded for a specific period of time. Putman and Walker (2010) considered goals to be one of three motivational components that need to be in place in order to maximize intrinsic motivation. The other two components are emotions and personal competency beliefs. The goal-setting process should be child-specific, based on his or her reading level, and should be realistic based on each student's circumstances and each student's availability for outside reading time as well as in-school reading time. Jenkins and Terjeson (2011) referred to this goal-setting process as the "Goldilocks dilemma" (p. 33). Teachers should set goals that are not too high or too low—but just right. By setting realistic goals, students should feel encouraged to read and meet those goals.

Beginning to use the AR program. When selecting a book, a student will check to ensure that the school has an AR test to go along with his or her choice. In this researcher's experience, if the book is not on the AR list, the student will return the book to the shelf and search until he or she finds a book for which a test has been purchased.

Librarians have made this much easier for students by labeling the books with the AR reading level.

The dilemma of choosing only AR books is not unique to this researcher's experience; it has been noted in several studies as an AR program disadvantage since teachers encourage adherence to reading only those books on the AR list (R. A. Johnson & Howard, 2003; Melton et al., 2004). The limitations on book choice can have a negative effect on student motivation to read. A study done by Esposito and Smith (2006) examined student motivation to read and found that students read more when their choices were not limited. Facemire (2000) looked at AR's effects on the reading comprehension of third-grade students and found that not only did students believe that their choices were limited but also, in many cases, parents would purchase books only if they could be tied to the AR program.

Additional studies, such as one completed by the FCRR (2006), although positive for AR, noted that a weakness of the program was that a major determining factor for students when choosing books was whether the school had access to the companion AR tests. Other researchers saw more long-term concerns with limiting students' choice of reading material. Facemire (2000) and Houston (2008) both believed that this restriction on choice failed to help students acquire a lifelong strategy for selecting appropriate literature.

Once the student has chosen and read an AR book, he or she may take a computer-generated quiz. The multiple-choice quiz is comprised of 5–20 questions depending on the difficulty level of the book. After the quiz is completed and submitted, the system gives the student immediate feedback by providing the score as a percentage

as well as the number correct. Students must score at least 80% to earn points; an average score of 90–100% on cumulative tests at a student's ZPD indicates proficiency at that level (Renaissance Learning, 2007). Teachers can access reports that supply the number of tests each student has taken, the books on which the student tested, the pass rate, and the number of points accrued. Teachers generally give the STAR assessment two to three times throughout the year to monitor progress and manage students' ZPD range.

Most of the process for students participating in AR is done independently, and that independence is not free from controversy. In this researcher's building, there have been several situations in which students have cheated by taking tests for other students or found ways to take tests more than once. Also, in this researcher's fifth grade class, students no longer have access to their user IDs; the teacher enters the code before each test is taken in an attempt to take away this temptation. This problem is not uncommon as others have reported similar concerns (Smith & Westberg, 2011; G. Thompson et al., 2008). A quick search on Google using the phrase "cheating with Accelerated Reader" generated several hits. One specific hit is a list posted on an Honors' English teacher's website (M. Thompson, 2012). This list includes a list of practices that constitute cheating, along with an AR contract that students and parents must sign. A further review of the list revealed numerous links to rules related to cheating on AR tests from schools at all levels as well as resources to help students cheat.

Program promises. The AR online Overview page makes several promises to teachers, one of which was examined in this review: "make essential reading practice effective for every student." The *AR Guide* (Renaissance Learning, 2007) provides the process through which AR states that the promise will be fulfilled. The heart of this

researcher's study is whether the extra reading practice time invested in reading AR books and taking the quizzes is effective as measured by student performance on the end-of-year SOL reading test.

The *AR Guide* (Renaissance Learning, 2007) encourages teachers to keep daily logs for student reading and test taking and to monitor book choices from both the school and classroom libraries. How does this close monitoring of 25+ students work in real life? It has been this researcher's experience that once the ZPD is determined, students' interactions with teachers are related only to their progress in earning points as shown through their quiz scores. This focus on score attainment was also articulated by Monroe (2005), who studied AR and aesthetic stance. Monroe indicated that teachers valued discussing point accumulation with students over more substantial interactions related to reading skill attainment. For this researcher, this dilemma of overemphasis on point accumulation has been consistent with teachers across six schools in two states. Teachers view AR as a time saver in that students are given their range for book selection and their movement toward their goal is loosely monitored. In the treatment group in this study, teachers indeed monitored students' progress at least every 2 weeks.

Synopsis of Opposing Pedagogical Views

Since the core of AR is the development of consistent reading practice, it is important to understand what this means from an AR standpoint. On page 5 of the *AR Guide* (Renaissance Learning, 2007), AR makes these statements regarding reading practice:

- AR practice is powerful practice.
- Practice must be on the right level of difficulty and cover a sufficient amount of time.
- Practice must be guided by the instructor.
- Practice must be enjoyable to sustain.

The *AR Guide* assures that AR enables powerful practice by

- Providing data that helps you monitor and personalize reading practice.
- Encouraging substantial amounts of practice, according to guidelines based on research findings.
- Making practice fun for students by facilitating successful encounters with text.

On the Renaissance Learning (2012) website, three types of reading practice for students are discussed: reading-to, reading-with, and reading independently. The practices of reading-to and reading-with students are related to using books that are at or above their ZPD level, and since these books may be difficult for some students, they may need to be paired with an adult or older student in order to navigate this more difficult text. AR's description of reading-with and reading-to involves the use of a person who is a stronger reader. However, Dixon-Krauss's (1995) study of partner reading stressed that teachers need to monitor these interactions for their pedagogical value rather than making the assumption that learning is taking place simply because one student is a stronger reader than the other.

Reading-with controversy. Other researchers, though, have a decidedly different concept of what reading-with means in good reading instruction. To them, these

processes involve more than reading the words in the text correctly. The teaching of reading is a deliberate process that involves teaching strategies to students and then monitoring the use of those strategies (Ferguson et al., 2011; Taylor, Pearson, Peterson, & Rodriguez, 2005). Guthrie et al. (2007) studied reading motivation and found that a high number of the students interviewed exhibited a strong, positive effect associated with the collaboration that comes from reading and discussing their reading with others. But pairing readers cannot be a random process. In their study that looked at the motivation for reading comprehension, Anmarkrud and Braten (2009) found that in order for the pairing of students to be a successful strategy, students must be given time to discuss not only the book they are currently reading but also its connection to the topic in general as well as its connection to other books read. This need to talk about what one has read was observed in a study examining reluctant readers conducted by Clark and Osborne (2007). They noted that students who are average to strong readers have a compelling need to discuss the books they have read, and simply pairing with another student to read is not sufficient to meet that need. A surprising as well as disappointing finding by Pavonetti et al. (2003) was that participating teachers did not allow students to share what they had read because of the fear that students would learn enough about the book from the discussion to take and pass the AR test. However, if a simple discussion between students provides enough information to pass an AR test, then a finding is confirmed from the studies by Cuddeback and Ceprano (2002) as well as Groce and Groce (2005). These researchers believed that the makeup of the questions on AR tests fosters lower-level comprehension of what has been read.

Although the *AR Guide* (Renaissance Learning, 2007) adds book talks to the strategy list, there is no discussion related to the true activating of prior knowledge or checking for prior knowledge before a student reads a book on which he or she will take a quiz. The AR definition describes a book talk as being more akin to a book commercial that advertises a particular book. Most studies reviewed noted the benefits of discussions with students prior to reading in order to activate prior knowledge regarding the text. While looking at three students and their interactions with the same text, Alexander and Jetton (2000) determined that the most influential factor in student comprehension of text was the activating of the prior knowledge they brought to the text. This absence of activating prior knowledge is seen as a huge loss for students. This loss was shown in Brozo's (2010) study on the "fourth-grade slump" (p. 148) in which he found that students at that level can usually read well as they possess good decoding and word attack skills, but their lack of relevant prior knowledge limits their comprehension of text.

Dzaldov and Peterson (2005) looked at prior knowledge from another direction: the book-leveling "mania" (p. 222). They found that the practice of leveling books in classrooms and libraries and in programs such as AR eliminated any opportunity to take student interest into account. The researchers recommended that teachers make certain they provide reading material that works to enrich and develop student background knowledge and experiences, not simply meet students' leveled-book needs. Glasswell and Ford (2011) also believed that teachers' professional judgment remains the critical factor when making book recommendations for students. Houston (2008) wrote that reading management programs such as AR do not help students develop lifelong

strategies for selection of appropriate reading material, nor do they create authentic recreational reading experiences.

Hilden and Pressley (2007) looked at teachers' ability to teach comprehension strategies to students and found that an area that lacked emphasis in instruction was that of triggering prior knowledge of a topic before reading. Taboada, Tonks, Wigfield, and Guthrie (2009), in their study on the variables that affect reading comprehension, asserted that utilizing prior knowledge along with allowing students to generate questions about what they have read are strong contributors to helping them build a "coherent mental text-representation" (p. 11).

AR's focus on the benefits of reading practice was discussed in a study by Topping, Samuel, and Paul (2007b). In their study, the researchers examined data from 45,670 students in Grades 1–12 that showed that these students had read over 3 million books. The conclusion of the study was that "practice does not make perfect—but attuned, successful practice makes perfect" (Topping et al., 2007b, p. 262). Simply providing time or requirements for reading will not increase student comprehension skills. Teachers' close, hands-on monitoring of student reading progress (not points gained on quizzes) will increase student reading comprehension.

Feedback as motivation? AR's goal is to stimulate students to read and to promote the enjoyment of reading. The process by which this stimulation and promotion is accomplished by the program is through the immediate feedback given when AR quizzes are taken (Renaissance Learning, 2007). Often, though, the promise of immediately knowing one's test score or the attainment of points is not enough to motivate a student to choose reading as a pleasurable activity. A selling point of AR is

that students will develop a desire to read as they experience success through passing quizzes and accruing points. In 1996, Carter wrote an article in which she described her observations related to AR and similar programs. She argued that reading as a pleasurable activity was devalued when rewards were given as motivation and that tangible rewards actually diminished student motivation to read. Motivation to read purposefully and to take meaning from what one reads requires more cognitive buy-in by a student. It is essential for teachers to learn that they have an important role in students' motivation to read and their intentional comprehension of what they have read. Alexander and Murphy (1998) considered this intentional interest in reading to be a motivational construct whereby students make a personal investment in their reading achievement.

When attempting to determine the motivating factors as to why some students read and others choose not to read, Dai and Wang (2007) examined reader need for cognition and reader beliefs and their relationship with text comprehension and interest development. Need for cognition (NFC) denotes cognitive effort as it relates to enjoyment of reading rather than as a deterrent to reading. Students who are high in NFC tended to obtain their information firsthand, and they used information gained through real-life interactions or interactions with text, both fiction and nonfiction, to shape their world (Dai & Wang, 2007). Students who were low in NFC tended to rely on others to gain information to support their opinions or decisions (Dai & Wang, 2007).

Dai and Wang (2007) as well as Schraw and Bruning (1999) looked at the implicit belief models that students bring to reading and how those models benefitted students' comprehension and use of text. The transmission beliefs model views reading as a

process that is a one-way, linear interaction between the student and the text; in other words, to this kind of reader, the text is unambiguous and unquestioned by the reader. In contrast, the transaction beliefs model is more dynamic and interactive; the reader processes and organizes information while reading and actually constructs new mental models based on the newly gained information. The transaction beliefs model is ultimately the goal of any teacher's reading program, yet it is still questionable as to whether these benefits can be gained from a program that assesses learning gained from reading a text through the use of a multiple-choice quiz after each reading.

In a study conducted by Law (2008), the researcher sought to determine if there was a relationship between what students believed about their own intelligence and ability and their comprehension of what they had read. The study was designed to look at whether students realized that they had developed reading strategies and whether they knew when to employ those strategies. The study found that when students were taught specific strategies to use while reading and they felt competent to use these strategies, they were more motivated to read. The result of the reading was also richer: Students were more engaged with the text and exhibited a stronger grasp of the content.

The success these students found increased their intrinsic motivation to read. But the simple knowledge that they possessed these strategies was not enough to cause the rise in achievement. Law (2008) found that a vital part of the success of the study was the instructional interventions provided by the teachers as students read. The teachers intervened as students read to determine which strategies they were using and to discuss why those strategies worked. The teachers also assisted as students attempted to figure out which strategy would work best when they developed problems with comprehension.

Law's (2008) description of a rich, student-centered instructional environment contrasted with the Guided Independent Reading setting described in the *AR Guide* (Renaissance Learning, 2007). The Guided Independent Reading setting in an AR classroom is seen as a time when students are at various stages of the reading/testing process. Some students are reading and some are preparing to take quizzes. During this time, the teacher is "circulating around the room, monitoring, coaching, and intervening" (Renaissance Learning, 2007, p. 6). If students require instructional interventions and strategy development in order to find success in reading and motivation to continue reading as Law found, it would be difficult to provide those essentials in such a setting.

This researcher's study supports the notion that independent reading is important. However, the quality of the reading time and the desire to continue reading is derived from the relationship between the student and the teacher and the use of instructional interventions that encourage strategy development during the reading practice. Taylor et al. (2005) determined that the most accomplished teachers were those who led students in discussions of the texts they were reading as they collectively constructed meaning from those texts. Ferguson et al. (2011) also found that the students who were most successful with reading were those who actively participated in their learning and discussions of their reading rather than taking on a passive role in the classroom. Both studies emphasized the importance of strategy instruction and teachers' purposeful involvement while students are practicing and learning.

Analysis of the Literature

There is no lack of studies that show success for those schools who use AR as part of their reading programs—many of which have been found to be little more than

anecdotal evidence without scientific merit. There is also an increasing number of researchers whose studies refute the results that show magnanimous gains in student comprehension, including Schacter (2000), whose meta-analysis of reading programs determined that there is little conclusive evidence that the AR program improves student reading comprehension or motivation to read. When thousands of instructional dollars are at stake, it is best to look at research that can be replicated and generalized across populations in order to make informed decisions (Blankstein, 2011; Carnine, 2000).

Following the path of the numerous studies on the AR website proved difficult when attempting to determine which ones were the initial studies. Albeit not an unusual practice in research circles, it seems that several of the authors had conducted original studies as teams, after which the individual authors used the same data in future publications. Although the same data and findings were used repeatedly, the AR website listed each as independent studies that confirmed their claims. As the review of the literature continued, this researcher became puzzled, not in regard to the statistical results of the many studies but at the studies that seemed to have been conducted in a manner that appeared more haphazard than controlled. Since these studies were being utilized to prove the claims of a program that requires an investment of several thousand dollars per year, the experimental management of the studies is under closer scrutiny than the reported results.

To assist school leaders through the maze of studies and articles, the WWC, a part of the USDE, was created to provide information related to the best research-based programs available for use with students. Many schools use this site since Title I funding requires that schools use those funds to purchase programs that have been approved

through federal scrutiny. The AR section on the August 2010 List for Adolescent Literacy provides a list of 318 studies, of which one was accepted—Bullock (2005)—while another—Nunnery and Ross (2007)—met standards, with reservation. In contrast, when searching for all studies related to AR, the Renaissance Learning site provides a list of 166 studies that are considered positive regarding the results of AR. This review pulled from both of those lists to provide a clearer picture of how AR has been studied in the past and why the current study was needed to add to the studies that provide both a treatment and a control group along with standardized data with which to compare comprehension growth.

The single study that was accepted by the WWC was the dissertation completed by Bullock (2005) that looked at how AR affects the reading performance of third-, fourth-, and fifth-grade students. Bullock's quasiexperimental study looked at three areas:

- Fluency as measured by the Dynamic Indicator of Basic Early Literacy Skills (DIBELS)–Oral Reading Fluency (DORF). The DORF was given as both the pre- and posttest for fluency.
- Comprehension was measured by the STAR Reading assessment through a pre- and posttest.
- Vocabulary was measured by the 4J Vocabulary Assessment, which consists of 90 vocabulary words that have been selected from a list of words found in *World Book*.

Bullock's (2005) study lasted 10 weeks and examined the reading performance of students who utilized the AR program and students who did not. Students in both groups were given 90 minutes of independent reading time each week along with 30 minutes of

library time per week. Students were tested via STAR Reading prior to the study to gain both a ZPD for their reading level and a pretest score. Students were tested via STAR Reading at the end of the 10-week study as well. Again, no significant main effect was found for comprehension between the treatment and control groups. Bullock found that the scores for DIBELS, the instrument used to measure oral reading fluency, showed no significant difference between the treatment and the control groups. When analyzing the 4J Vocabulary Assessment, Bullock was able to use data only from the fourth-grade students due to errors in data collection for Grades 3 and 5. But, again, no significant difference was found between the treatment and control groups. Consequently, Bullock found that the results of his study did not support the claims made by Renaissance Learning on its website.

Bullock (2005) stated that the first limitation of his study was the duration, as it lasted for only 10 weeks, which limited students' exposure to AR to only that period of time. The number of participants was also low; no classroom had more than 12 students. He was concerned as well about the level of buy-in by teachers, which could have affected the outcome. He also believed that the nonrandomization of students to the treatment group was a limitation. He admitted that these limitations are a part of the negatives of attempting research in the public school setting and believed he had controlled for as many of these confounding variables as possible.

When the search for AR began, several names consistently appeared. A number of articles were either authored or coauthored by Keith Topping, a professor of educational and social research at the University of Dundee, Scotland. According to the University of Dundee (2012) website, Topping is a member of the International Reading Association

task force. A major purpose of the task force is to disseminate the data from research studies related to Program for International Student Assessment and Progress in International Reading Literacy achievement assessments used in 37 countries. Topping has written several articles connected to studies he either conducted on the effectiveness of AR or on his analysis of previously collected data. His studies are generally positive toward AR, but his objectivity has been questioned (Biggers, 2001) due to the fact that he has cowritten at least four studies with the creator of AR, Terrence Paul (e.g., Topping & Paul, 1999; Topping et al., 2007a, 2007b, 2008) and wrote the foreword in Paul's (1996) study, *Patterns of Reading Practice*.

In 1996, Paul conducted a study on student reading practice. Topping wrote the foreword and asserted that "those schools which have implemented the Accelerated Reader daily progress-monitoring assessment tool for longer show higher rates of reading practice" (p. 5). Thirteen thousand U.S. schools that had owned AR for more than 4 months were surveyed. Paul's study looked at survey data on 659,214 students in Grades K–12 from 2,193 of the responding schools for one school year. His data indicated that, again, the peak AR years are Grades 3–5. The survey did not collect individual student point data, so data from his 1992 and 1993 surveys were analyzed again to show the variation in reading practice for both low- and high-performing students. Even though the survey data were from 1996 and the point data were from different studies in 1992 and 1993, this did not prevent Paul from stating that the correlation found was that the larger the number of points accumulated by students, the higher their 1994 National Assessment of Educational Progress scores. But the question still arises as to whether this result confirms the Matthew Effect as related to reading: that better readers have

better scores. As with other studies, the WWC (2010) rejected Paul's study due to the lack of a control group for comparison.

The first Topping study to be examined was conducted with Vollands and Evans (Vollands, Topping, & Evans, 1999). The study was a quasiexperimental action research evaluation that looked at AR and its formative effects on both reading achievement and motivation to read. Two elementary schools in Aberdeen, Scotland, were involved in the study, both with high populations of disadvantaged students. The schools were chosen because of the availability of computers as well as the schools' willingness to participate in the study. Both schools had a treatment and control group and both groups were comprised of a majority of students with reading difficulties. The first school's treatment group consisted of 27 students, while the control group (mixed age/mixed ability) contained 12 students. The study began with students receiving 15 minutes per day of reading time, which increased to 30 minutes after 5 weeks. An additional 30 minutes of time was provided for the teacher to read to the students. Students were also able to test on teacher-read books, and did so with a high rate of success. The Discussion section of the study brought out several factors that affected the results. There was a noticeable lack of books for students at lower reading levels. The library's collection of books were not consistently marked with the correct AR level, and students paid little attention to those books that were leveled as they consistently chose books from outside of their ZPD. Although the school was chosen based partly on its access to computers, a shortage of computers on which students could test developed as the study progressed.

The second school in Vollands et al.'s (1999) study compared 11-year-old sixth-grade students (24) with 10-year-old fifth-grade students (26). It was noted that fidelity

of implementation by classroom teachers appeared only near the end of the study. Reading time also increased during the study from 15 minutes to 20–30 minutes. The school also had difficulty providing appropriately leveled reading material for students and had difficulty with computer access.

In both schools, the researchers had difficulty accessing the necessary resources to provide adequate access for the experimental group. Computer access for testing was limited and there were only a small number of AR books available for students. Despite the admitted negatives related to the confidence in the study, the researchers asserted that gains were "superior to gains from regular classroom teaching and an alternative, intensive method" (Vollands et al., 1999, p. 197). The statistical results indicated that AR students showed gains (110.08 vs. 117.27) as indicated by their performance on the Neale Analysis of Reading Ability (Neale, 1989) in the area of reading accuracy, whereas the control group's scores stayed virtually the same (109.83 vs. 109.75). However, the chart on page 205 of the study illustrates that when students' scores were analyzed to determine their performance in reading comprehension, the AR students' scores decreased slightly (105.33 vs. 104.64), while the comparison group's scores went down significantly (96.42 vs. 84.58).

Even though Vollands et al.'s (1999) study clearly contained questionable processes and results, it did not keep AR from listing it twice in its Literature section, which contains studies that confirm its claims of success with the program. In contrast, according to the 2010 WWC report, the results of the study were inconclusive due to the fact that any improvement in student performance could not be directly related to the intervention (AR).

Topping conducted another study, with Sanders (2000), which looked at the archived data from the Tennessee Value-Added Assessment System (TVAAS). The TVAAS includes student achievement data from all Tennessee Grade 2–8 students in public schools. The AR data for 62,739 Tennessee Grade 2–8 students (provided by Renaissance Learning) were merged with the TVAAS teacher effects database, and the study examined the relationship between these two pools of data. The comparison achievement scores entered into the TVAAS system were those of the Tennessee Comprehensive Assessment Program (TCAP), a test that must be taken by all students in Grades 2–8 at the end of each school year. The TCAP scores were converted to grade equivalents in order to be compared to AR grade equivalents. The purpose of the study was to determine the value added to teacher effectiveness through the use of a program that monitored student reading habits as well as their performance on quizzes related to the books they read over time, as well as to measure teacher effectiveness with the program.

The Discussion section of Topping and Sanders's (2000) study brought out many concerns as to the ability to trace effects back to the intervention. Selection bias may have been an issue with the study as the researchers experienced 21% sample attrition with AR records. It was believed that attrition may have had an effect on the ability to generalize the results. Difficulties also arose when attempting to follow the student scores to the teacher/classroom level to identify the teacher effects related to the study. Findings showed that students across all ability levels tended to read books below their actual grade level. The study also found that the value-added benefit tended to rise when student scores were higher on the quizzes. Results indicated that middle-level readers

(average readers) tended to benefit the most from AR—with the most value being added when students read more difficult texts.

As with several of the studies done with data gathered from large populations, there is little direct knowledge of the manner in which the program was implemented. There is no information regarding whether confounding reading interventions were introduced along with AR and no monitoring of students' book selection habits or of their test taking. This lack of attachment to the sites limits the ability to use the data as definitive in the study's assertion that AR use was the cause of any increase in the TCAP scores. This point was brought out in the study as the researchers attempted to determine if teacher implementation of AR had an effect on any improvement in reading comprehension (Topping & Sanders, 2000). The concern was how to suitably establish if any increase in skills could be related to teacher behavior. The behaviors in question were whether teachers encouraged students to read more difficult AR books or whether the increase was a result of the use of more difficult stories from the basal reader or content-area texts. No definitive answers were to be found in the study's results.

The results of Topping and Sanders's (2000) study, which showed that students who read more AR books also had higher TCAP scores, bring up the question of the appearance of the Matthew Effect with readers. This effect is mentioned in several studies either directly or through similar descriptions of the outcomes when reviewing the results of AR (Borman & Dowling, 2004; Dai & Wang, 2007; Guthrie et al., 2007; Hansen et al., 2009; D. W. Johnson & Johnson, 1999; R. A. Johnson & Howard, 2003; Palmiter, Arcaira, White, & Reisner, 2009; Pappas, Skinner, & Skinner, 2010; Retelsdorf, Koller, & Moller, 2011). The Matthew Effect refers to the paraphrasing of the biblical

passage in Matthew 25:29: The rich get richer and the poor get poorer. Initially, this circumstance was associated with economic situations studied in sociology, but the concept has been adopted by researchers such as Stanovich (1986) and Hansen et al. (2009) to describe students' chances of becoming competent readers when they begin school with deficit reading readiness skills. The dilemma encountered when examining a study in which average readers or above-average readers seem to benefit the most is determining whether the benefit is the result of the Matthew Effect, in which good readers tend to get better regardless of the intervention introduced. Although AR used Topping and Sanders's study as proof of its effect on student reading performance, when the WWC reviewed the study, it was found to be ineligible for use due to the fact that no control group had been utilized.

Topping (2006) conducted a study in several Specialist Schools in Greater London as the schools were implementing Accelerated Math and Accelerated Reading. Their primary interest in conducting the study was student performance in reading. No control groups were utilized, and although the results showed statistically significant gains, it was still evident that students who read well often performed much better than those who did not. The median score gain for those who read more than 20 minutes per day was 117.49 compared to 77.17 for those who read less than 20 minutes per day. However, even students who read more each day showed an average AR test score of below 85%. The goal of AR is to have students consistently scoring between 95% and 100% on their AR tests. Technically, students did not show success as AR test scores were below the percentage that is considered mastery of reading material at their ZPD level (Topping, 2006).

Topping continued to study the effects of the AR program. In 2007, Topping et al. used the same research data to produce three studies (the third was published in 2008). The first study (2007b) attempted to determine if reading practice made for perfect practice. The researchers used the number of points earned through AR testing by each student as a measure of time spent on reading. The study was considered to be correlational rather than experimental as the results indicated that quality of practice was considered to be higher when students passed a greater number of tests and earned a higher number of points. The study ends with the statement that using stored data to determine student practices is much more cost effective than traditional methods of collecting data at the school level. This use of stored data, however, eliminates the possibility of observing firsthand whether programs are implemented effectively as opposed to drawing conclusions from a data set.

When these same data were used a second time, Topping et al. (2007a) explained that they believed that part of the difficulty with their 1999 and 2000 studies, mentioned earlier, was the fidelity of implementation of the AR program expectations. The AR expectation to be studied was that of providing students with 30–60 minutes per day of quality reading time. The researchers used the data for this second study to look more closely at the impact of implementation differences. Again, the study looked at the same archived AR data. Although data collected were from students in Grades 1–12, the majority of the data were from Grades 1–6. Of the participating schools, 43% of the teachers had achieved some level of certification, indicating that they had met benchmark standards for minimum quality implementation of the AR program. The authors added, however, that this certification did not guarantee that teachers still practiced high-quality

implementation. Interestingly, the majority of the school respondents for the study were in their first or second year of AR use. Those schools that had used AR for longer periods of time declined to participate.

A noteworthy finding of the second study (Topping et al., 2007a) was that as the grade level increased, the implementation quality index decreased. The study was also one of several that showed that performance on and participation in AR is much higher in the lower grades (Hansen et al., 2009; Nunnery & Ross, 2007; Nunnery et al., 2006; White & Reisner, 2007). Once more, the WWC (2010) rejected the study due to the fact that the researchers provided no comparison data from a control group.

In 2008, Topping et al. published another study utilizing the same archived data as the 2007a and 2007b studies. The purpose of the third study was to examine the relationship of challenge, nonfiction, and gender to student achievement. Although the study does not indicate that AR had any effect on student performance, it is listed on the 2011 AR website as a study that supports AR's success statements. Conversely, the use of AR in the study was actually discussed as a limitation due to the fact that the only texts examined were those for which an AR test was available. Topping et al. believed the results could not be generalized to other non-AR situations.

Samuels conducted studies on AR with other researchers besides Topping and Paul. In 2003, he joined Lewis, Wu, Reininger, and Murphy to compare classrooms that utilized the AR program with those classrooms that did not utilize AR. The quasiexperimental study was very closely monitored and included tightly controlled conditions. Teachers were randomly assigned to treatments, and reading time in both groups was consistent throughout the study. The study lasted for 6 months, and pre- and

posttest data were collected from all students. The STAR Reading test was taken by students in both the control and experimental groups in order to provide an AR reading level for each student. STAR was also used as a posttest to measure growth or lack thereof in student reading levels. Students also took a Curriculum-Based Measurement test to measure fluency at the start of the study and at the end. The results showed that the students who used AR scored higher in comprehension and vocabulary in both grade levels.

Although Samuels et al.'s (2003) study was tightly controlled, the number of participants was quite low—the control group included 19 students in third grade and 20 in fifth grade, while the experimental group included 16 students in third grade and 19 in fifth grade. The WWC (2010) rejected the study due to the changes in the participants. The attrition rate of the students during the study, from 74 to 67, exceeded the WWC standard, and it was believed that this rate compromised the integrity of the data.

Samuels and Wu (2003) utilized the aforementioned data in another study to determine the effects of immediate feedback on student reading achievement. This second study added another variable that was not mentioned in the first study: The control group of students did not take AR tests, but they were required to write a book report after each book read. Students who are required to spend the time to produce a book report after each book read will, simply by the nature of the time required, have less time for reading. This one variable could have had a significant effect on the results of the study. However, according to Samuels and Wu, the results remained positive for the effects of the immediate feedback given students at the end of an AR quiz. The authors asked the question as to whether the results of the study proved the Matthew Effect since

the results showed that high-ability readers took and passed more tests than low-ability readers. This researcher would question the wisdom of believing that producing a book report and taking a 10-minute quiz after reading a book are comparable activities for students. As with the first study, the WWC (2010) rejected this second study due to the attrition rate of participants, yet AR posts this study as a support for adding AR to schools' instructional programs (Renaissance Learning, 2012).

Samuels and Wu (2004) used these data again for a paper presented at the 2004 International Reading Association Convention. The paper focused on the relationship between the amount of time spent on reading independently and its effects on achievement in reading. Although the study showed positive reading achievement results for students in the AR group, it must be remembered that while the AR group was reading, the control group was writing book reports. The WWC (2010) did not review this study as it had already rejected the initial study due to the high attrition rate of participants.

Nunnery et al. (2006) conducted a study on 978 students in Grades 3–6. The students in the study utilized the AR program and the teachers participated in Reading Renaissance, a Renaissance Learning product that is designed to train teachers in the proper use of the AR components: reading-with, reading-to, and reading independently (Renaissance Learning, 2012). Participating schools volunteered for the study, which required principal permission and at least two teachers at each site. Teachers at each school were randomly assigned to the control or treatment group. Of the initial 1,023 students involved, only 978 students had a pre- or posttest score—97 were missing the

pretest scores, 176 were missing the midterm scores, and 146 were missing posttest scores.

The results of the study supported AR as a tool that helped increase student performance in reading, with effect sizes higher in third grade than fourth and fifth grades (Nunnery et al., 2006). This difference may have been the result of observations made by consultants that noted that adherence to established reading times in the experimental classrooms was much lower in fourth and fifth grades. Teachers at those grade levels also experienced more technical problems with the AR computer program. One finding of the study contradicted that of Topping et al. (2007a, 2007b), who found that fidelity of implementation was crucial to increased achievement for students. Nunnery et al. (2006) felt that the quality of the implementation of the program (which was lacking in the study, even with the addition of Reading Renaissance) did not predict student achievement growth. They also determined, as this researcher did, that although there are many published articles and studies related to the effectiveness of AR, most of this evidence is based on anecdotal reports based on increased test scores, correlational data, and studies that did not use control or comparison groups.

Nunnery et al.'s (2006) study revealed a pattern that began to appear as studies were reviewed: that participation in AR and achievement with AR tend to peak in third and fourth grade (Harper, 2001; Kambarian, 2001; Ross, Nunnery, & Goldfeder, 2004; Topping & Paul, 1999; Walberg, 2001; White & Reisner, 2007). It seems that although the AR program is advertised to serve K–12 students, there is little documented participation after fifth grade. This pattern has been constant across schools in this researcher's experience as well. The WWC did not accept the study due to the high

attrition rate as it was believed the researchers kept the comparison groups from being equivalent, but Nunnery and Ross conducted a similar study in 2007 that was accepted with reservation by the WWC (2010).

The Nunnery and Ross (2007) study that was accepted by the WWC with reservation was accepted based on effectiveness ratings on the findings from certain grade levels. The WWC accepted results gained from a cohort of 912 fifth-grade students divided into a control and intervention group and a second cohort of 891 fifth graders. The researchers used repeated-measures analysis of covariance (ANCOVA) on transformed variables to test the effects of the program using Texas Assessment of Academic Skills data for comparison. The first cohort's progress was reviewed after 2 years, and it was found that the program type had a statistically significant effect as shown by the covariate-adjusted means on the averaged transformed variables. The 1999–2000 and 2001–2002 scores were averaged together; the control group $M = 1.04$ ($SD = 0.33$) and the experimental group $M = 1.11$ ($SD = 0.32$). The effect size estimate, Cohen's *d*, was 0.22.

The progress of the second cohort was reviewed after 3 years (Nunnery & Ross, 2007). The covariate-adjusted means on the averaged transformed variables for the control group were $M = 1.10$ ($SD = 0.34$) and for the experimental group (AR) $M = 1.15$ ($SD = 0.35$). The researchers believed the fifth-grade scores showed more growth using AR due to the fact that students had more years of experience with the program. The researchers also noted that the findings would have had more strength if the populations studied had been more randomized in their assignment to the control or intervention group—a feat they admitted is quite difficult in a public school setting.

Although Nunnery and Ross found positive results for the use of AR in their 2007 study (with fifth-grade students, but not eighth-grade students), they had concerns over the population used. In their Discussion section, they noted that the results may not generalize across other populations due to the fact that the schools involved were relatively high-performing suburban schools.

Some studies that show support or marginal support for AR raise the specter of lack of objectivity. Husman et al. (2005) completed a yearlong study on student goal orientation and assessments using AR test information. Their study was conducted with a grant from Renaissance Learning, the parent company of AR. The impetus for the study was the concern voiced by parents and teachers that frequent testing could negatively impact student goal orientation and performance. The results showed that student mastery orientation remained strong throughout the year, with only a slight decrease in performance orientation. The Discussion section of the study brought out a positive finding that this researcher believes is at the heart of any student performance concern; this same relationship between a teacher's characteristics and his or her students was also revealed in a study conducted by Taylor et al. (2005). The study's results indicated that frequent testing with AR did not have a negative effect on students' continued desire to set and meet goals. But since the researchers did not utilize a control group for comparison, the study was rejected by the WWC.

The aforementioned school was used in a study of longer duration by the same researchers (Brem, Husman, & Duggan, 2005). The study continued for another 2 years with the addition of a control group that used AR but had not adopted the Reading Renaissance teacher training program. The study employed multiple data sources for

measurement: SAT9; demographic information; STAR Reading tests; AR data; survey data from parents, teachers, and students; focus groups; reading specialist and librarian interviews; and student data from the Patterns of Adaptive Learning Strategies. Although a strong correlation was seen between STAR scores and SAT9, an unexplained variance (47–56%) in STAR scores led the authors to place more trust in the SAT9 scores for a true measure. They believed the STAR could be used as a predictor of SAT9 performance.

Again, the researchers found that students' desire to perform well on the AR tests did not decrease with frequent testing (Brem et al., 2005). A disappointing finding was that students involved in AR could not differentiate between trying to develop better reading skills and accumulating more points—a trait that is tied to student performance orientation (why they want to do well). Student performance orientation was low at the onset of the study and decreased throughout. It seems that students did not have a clear understanding of why they should perform well on AR tests. The researchers also determined that teachers were too focused on passing AR tests and, therefore, did not increase the difficulty of books recommended to students. The use of AR also produced a negative, competitive environment. This second study was not reviewed by the WWC, but it was listed as confirming AR's claims on the Renaissance Learning (2012) website. There is still a question as to the objectivity of the study due to the fact that it was conducted using a grant from Renaissance Learning.

Brem teamed with Sadusky (Sadusky & Brem, 2002) to use the same data on the schools in the two aforementioned studies with the intention of determining the effects of AR and Reading Renaissance on school-wide improvement. The results showed that

School A showed significant improvement in SAT9 scores over School B, with the implementation of Reading Renaissance given as the reason for the increase. However, the researchers disclosed that School A had implemented several strategies along with Reading Renaissance, which may have impacted the positive rise in scores; School B was described as utilizing the AR program sporadically, even "haphazardly" (Sadusky & Brem, 2002, p. 5). The concern with accepting the validity and reliability of the data is that the second school had embraced a different type of reading program that included tests and points, and students rarely used AR. As the use of the same data for various studies is reviewed, it seems that further explanation as to the actual use of AR, as well as other, possibly confounding, interventions surfaces. This lack of complete information regarding the true variables in each study causes the reader to pause and reflect on the accuracy of the findings. These bona fide concerns related to the integrity of the information did not prevent AR from listing all of the studies on its website as evidence of positive effect. And, as mentioned in the previous studies, funding was provided by Renaissance Learning. The WWC (2010) rejected this study because it did not consider School B a comparison group.

Other studies used to support AR's claims that were rejected by the WWC due to the lack of a control group include McKnight's (1992) qualitative study examining students' attitudes toward reading after using AR. She looked at the TV viewing habits of fifth-grade students as well as their book checkout habits in the school library and their reading logs. Seventeen students were studied. Only 10 students kept accurate reading logs; five others kept logs but were inconsistent in recording their activities. Classroom observation of sustained reading time showed that 14 of the 17 students maintained

appropriate reading habits during the reading time. Although the researcher's goal was for seven of the 17 students to choose reading as their second favorite activity, only two rated reading as second, with 13 rating TV watching as their favorite activity. Some students increased their reading, but they replaced other activities with reading—that is, they did not change their TV-watching habits. Despite the low number of participants and the minuscule changes in behavior, the author contended that her results showed that students' attitudes improved based on responses on pre- and postsurveys related to TV watching. Her study is listed on the Renaissance Learning (2012) website as evidence for AR's positive effect on student reading performance and motivation to read.

Goodman (1999) also saw positive results in his study of 282 middle school students who used AR. His study used two forms of the Gates-McGinitie Reading Test as the pre- and posttest as the measure of progress. According to the measure, the students showed significant gains in reading. Although the study showed significant gains in vocabulary and vocabulary and comprehension combined, there were no statistically significant gains in grade-level equivalency comprehension. However, since Goodman did not use a comparison group, it is difficult to ascertain if the growth was natural growth gained from a year's schooling or truly the result of the students' experience with AR.

Rodriguez (2007) looked at student achievement on the California Standards Test (CST) for reading. He looked at eighth-grade AR data and compared it to the end-of-year test results and found that the higher the AR participation, the higher the score on the CST. A closer look at his results indicated that positive results were indicated for students who had amassed 100 points or more during the study. The study certainly

supports the concept of the Matthew Effect with these readers. If the research identifies AR as a program that keeps good readers reading, what also surfaces regarding those marginal readers who reluctantly choose reading as a pastime? Rudd and Ware (2006) examined the effectiveness of reading practice and found that when using AR, students made significant gains in reading over the yearlong study period. What the researchers also found was that the percentage of students who said they read every day decreased by 8 percentage points and that, in the end, students still preferred watching TV to reading. Souto-Manning (2010) looked closely at AR from the student perspective and found that, although many students participated in AR, they did not like the program.

The studies reviewed in this literature review were chosen due to the immense power they have been given through the Renaissance Learning (2012) website. Each study was listed as showcasing the positive effects AR can have on student progress, yet when they are placed under scrutiny, the setting, participants, and even results can be seen as questionable. This lack of trust in the advertising of the program was the driving force behind this researcher's study. It is hoped that by providing current data using an AR group compared to a non-AR group, other researchers will conduct their own studies in order to look more closely at programs such as AR to measure their true effectiveness across other populations.

Gender Differences in Reading

One aspect of this study involves examining any differences that may surface related to the reading performance of males in comparison to females. It has been this researcher's experience that females outperform males on reading tasks in elementary school. Although this phenomenon has been the topic of many grade level meetings with

staff, these meetings have rarely included discussions of research related to the concern. In reviewing the literature, it was found that there are several possibilities as to why male students tend to read at a lower level than females. A study conducted by Below, Skinner, Fearrington, and Sorrell (2010) measured 1,218 students, k-fifth grade, on four measures: phonemic awareness, phonics, orthography, and fluency. Initially, they found that male students began kindergarten with a deficit in Initial Sound Fluency (ISF). Male students entered kindergarten with average ISF scores of 8.68 whereas female students began the year with ISF scores of 11.40. Although both groups showed improvement with the mid-year assessment, the scores of male students still trailed that of females with males scoring 19.31 while females scored 23.72.

Although the scores for males and females showed a deficit for males at the beginning of kindergarten, Below, et al. (2010) found that, when examining the groups over time, males gained ground in reading and by fifth grade, the scores for Oral Reading Fluency (which includes a measure for comprehension) for males and females were nearly identical. When analyzing the possible reasons for the gain in reading, the researchers indicated that they felt their study supported earlier studies that indicated that females enter school with stronger literacy skills and that these skills increase throughout their school experience. However, their results did not support the notion that males do not catch up before the end of elementary school.

Below, et al. (2010) also felt that their study may also support the studies that have indicated that male students may not be as motivated as female students to read the type of text that is seen in the earlier grades as these texts may appeal more to female students. When students enter the fourth and fifth grades, they have more freedom in

determining the types of books that they read and males may feel more empowered to read books that are a best fit for their interests.

Logan and Johnston (2009) also looked at gender differences in both ability and attitudes. They found that females have a more positive attitude toward reading as a chosen recreational activity. The study showed that although attitudes toward reading for both genders tend to become more negative over time, females' attitudes tended to be more stable than males. An interesting finding in this study contrasted with the views that this researcher has had about possible reasons for the poor performance of males in reading. It has been believed that when males read poorly, it affects their attitudes about reading. This study found the opposite—if males have a negative attitude toward reading, that attitude will affect their reading proficiency.

In a later study by Logan and Johnston (2010), they explored the notion that gender should be considered a context in reading instruction. Instructionally, gender should be considered when teachers are planning their lessons. The study indicated that when studied through neuro-imaging, the brains of males and females behaved differently during tasks involving language such as reading. The study of the brain activity may signify that the male brain does not learn to integrate visual and phonological information simultaneously, a behavior that is required for the comprehension of what is being read. Males who are learning to read tend to be more apt to decode words rather than using pictures or context clues which are strategies that add to comprehension of text. These are skills that females tend to use more successfully at an earlier age than males. Instructionally, teachers should include lessons that are more intentional in helping male students develop those stronger comprehension strategies that involve skill integration.

An additional consideration when examining any academic concerns for males is that of discipline and problem behaviors. McIntosh, Reinke, Kelm, and Sadler (2013) studied the connections between reading proficiency, problem behaviors and gender in elementary school students. They first found that males had more office referrals for behavior than females. Their study suggested that problem behaviors may not be predictors of poor reading ability. However, since the types of problem behaviors exhibited by males tended to be more aggressive and disruptive, they have the potential to interfere with the continuity of reading instruction. Female problem behaviors tended to be more covert and less disruptive of the classroom setting. Their instructional recommendations included identifying the male students with problem behaviors early in their school careers and provide supports for intentionally changing those behaviors.

Theoretical Framework

The theoretical framework behind the AR program is that of Vygotsky's ZPD (Knapp, 2008; Kozulin, 2002; Kravtsova, 2009; Niewolny & Wilson, 2009; Vygotsky, 1978), which is a part of his sociocultural learning theory. The AR program's use of the ZPD refers to the range assigned to each student after he or she has taken a STAR assessment. This range (reading level) is then used by the teacher to slowly increase the student's reading level through the choice of specific books at the student's level and the successful completion and passing of quizzes attached to those books (Renaissance Learning, 2007).

When Vygotsky's (1978) sociocultural learning theory is examined, it reveals that his intent was to develop a theory through which students use problem solving, language, and social interactions to learn. Previously mentioned research showed that students who

learn problem-solving strategies and have discussions with teachers and peers related to their reading are more motivated to read and comprehend more of what they read. It is questionable as to whether AR program use fulfills the spirit of Vygotsky's ZPD.

Summary

This literature review revealed the many articles and studies available that have dealt with the effectiveness of AR on students' reading comprehension and motivation to read. The dilemma identified in this review is that of the researcher's ability to place trust in the results of those studies that support AR's claims of superior gains in reading scores for students through STAR or other standardized measures. Resources such as the WWC have been established to assist in the validation of results by providing reviews of the literature. The void in the literature is that of current studies that utilized both a control and experimental group of students in which a defined group of students used the AR program consistently. This study was designed to add to the body of literature that includes quasiexperimental evidence utilizing a control group that utilized a similar reading program through the system's literacy plan while the treatment group of fifth-grade students utilized the system's literacy plan along with the mandatory use of AR throughout the year.

CHAPTER THREE: METHODOLOGY

The purpose of this causal-comparative research study was to add to the body of literature that has been produced examining the effectives of the Renaissance Learning product Accelerated Reader. The design is that of a causal-comparative, nonrandomized control group with a pretest and posttest. This design works well within the school setting as classes cannot be organized to accommodate research during the school year (Ary, Jacobs, Razavieh, & Sorensen, 2006). The study utilized end-of-year SOL reading test data for the classes from two groups of fifth-grade students. Archived data from fourth grade were compared to the Spring 2012 fifth-grade test data.

The results of this research study provide school administrators and central office administrators with current information to assist in determining if AR is worth the considerable financial investment required to keep it in schools. There are noticeable discrepancies in the literature related to whether the process through which students read books and then complete AR quizzes on those books truly provides effective practice that improves comprehension (Cuddeback & Ceprano, 2002; Esposito & Smith, 2006; FCRR, 2006; Houston, 2008; Krashen, 2005; Macalister, 2011; Melton et al., 2004; Nunnery & Ross, 2007; Pavonetti et al., 2003; Schmidt, 2008; G. Thompson et al., 2008).

Research Design

A causal-comparative, or ex post facto, study was used to determine if the reading test scores of randomly chosen students from two intact groups were affected by one group's use of AR. Because students could not be assigned to the experimental or control group, randomization was accomplished by selecting students anonymously from the mandated AR group list and the non-AR group list. A causal-comparative study, or ex

post facto design, can be described as a study conducted in reverse. Where experimental research randomly assigns groups to two or more conditions that are identical, with the exception of the application of the independent variable, a causal-comparative study examines different intact groups and attempts to determine the conditions that occurred as a result of that difference (Gall, Borg, & Gall, 1996). Although a causal-comparative study lacks the level of scientific confidence placed in the results gained from experimental research, it can provide real-life compromises in conditions where it is not practical or not possible to conduct experimental research (Abrami, Cholmsky, & Gordon, 2001). Due to the inherent absence of control in a causal-comparative study, its results can sometimes be dismissed as predictors of causality; however, these results can be quite beneficial as they can raise questions or uncover relationships that can be the impetus to guide future experimental research studies (Ary et al., 2006).

This study examined the difference between the Virginia SOL reading test scores of two groups of fifth-grade students. The first group consistently used the AR program for a full year along with the system-required literacy plan, while the second group utilized only the literacy plan. Fourth-grade 2011 Virginia SOL reading test scores were used as the pretest and fifth-grade 2012 SOL reading test scores were used as the posttest. The SOL for reading under which fourth- and fifth-grade students have been instructed contain identical descriptions of the reading comprehension skills that are expected to be mastered and tested on the SOL test. The Virginia Department of Education (2012b) indicated that Standard 4.5 (h, k, and l) and Standard 5.5 (i, l, and m) for reading fiction require that students draw conclusions and make inferences from text, use reading strategies throughout the reading process to monitor comprehension, and read with

fluency and accuracy. Standard 4.6 (k and j) and Standard 5.6 (l and m) for reading nonfiction state that students should use reading strategies throughout the reading process to monitor comprehension and read with fluency and accuracy. The manner in which these skills are tested requires that students read passages and answer comprehension questions.

The dependent variable examined in this study was that of SOL reading scores based on the introduction of the independent variable, the use of the AR program. Sixteen elementary schools were used to provide the student pool for this study. Students for the experimental group (Group A) were randomly chosen from the schools that require mandatory use of the AR program for the entire fifth-grade school year. The participants in the control group (Group B) were drawn randomly from the remaining schools in the system. The demographic data for both groups are shown in Table 1.

Table 1

Student Demographics

Category	Group A (AR) % (n = 51)	Group B (non-AR) % (n = 57)
Male	49.0	57.9
Female	51.0	42.1
Black	20.4	10.0
White	66.7	81.7
Asian	5.6	1.7
Hispanic	3.7	6.7
American Indian	4.0	7.0
Special education	16.7	15.0
Disadvantaged	12.5	9.2

The first group utilized the AR program as a part of their reading instruction in addition to the system-required literacy plan. The second group utilized only the system-required literacy program. Although the literacy plan includes a reading curriculum that utilizes the adopted basal reader, it can also include the use of novels at teachers' discretion. The literacy plan also requires the use of a word study program as well as the use of a running record system designed by Fountas and Pinnell (Heinemann, 2010).

Treatments

Literacy plan treatment. The school system in which this study took place has designed a consistent literacy plan for all elementary schools. The plan includes

consistent requirements for the teaching of writing, word study, and comprehension and fluency. Although specific in measurement procedures, it is broad in that schools have autonomy in choosing instructional methods that meet the needs of their particular populations. All schools must have a word study component for students at each grade level. They must also measure reading performance according to the levels shown in Fountas and Pinnell (Heinemann, 2010). Both the control and experimental groups in this study followed the system's literacy plan, which has been in place for 2 years.

Accelerated Reader. The schools that mandate the use of the AR program with their fifth-grade students had similar participation requirements. The teachers required that students take the STAR assessment at the beginning of the year and periodically throughout the year. The STAR is retaken when students begin to perform consistently near the 100%-correct level on their AR tests or their scores are consistently low. AR schools also set point goal requirements based on the initial STAR assessment and adjust that goal accordingly throughout the year. Teachers monitored students biweekly to ensure they were on track to meet their goals. Grades could be affected if goals are not met.

Research Questions and Hypotheses

Research Question 1: What is the difference between the 2012 Virginia Standards of Learning reading scores in fifth-grade students who used the system literacy plan along with mandatory use of Accelerated Reader as compared to the scores of fifth-grade students who used only the literacy plan?

Null Hypothesis 1: There will be no statistically significant difference between the scores of fifth-grade students who used the system literacy plan along with mandatory

use of Accelerated Reader when compared to the scores of fifth-grade students who used the literacy plan without mandatory use of AR as evidenced by their performance on the 2012 fifth-grade Standards of Learning reading test.

Research Question 2: What is the difference between the 2012 Virginia Standards of Learning reading scores by gender for fifth-grade students who used the system literacy plan along with mandatory use of Accelerated Reader as compared to the scores of fifth-grade students who used only the literacy plan?

Null Hypothesis 2: There will be no statistically significant difference between the scores of female fifth-grade students who used the system literacy plan along with mandatory use of AR when compared to the scores of female fifth-grade students who used the literacy plan without mandatory use of AR as evidenced by their performance on the 2012 fifth-grade Standards of Learning reading test.

Null Hypothesis 3: There will be no statistically significant difference between the scores of male fifth-grade students who used the system literacy plan along with mandatory use of AR when compared to the scores of male fifth-grade students who used the literacy plan without mandatory use of AR as evidenced by their performance on the 2012 fifth-grade Standards of Learning reading test.

Data Collection and Analysis Methods

Data collection for this causal-comparative research study was completed through accessing archived test scores for both groups to obtain students' fourth-grade SOL reading test scores. This same process was used in collecting archived 2012 fifth-grade reading test scores. This researcher gathered scores by student name; a research assistant was used to assign anonymous identifiers to the test scores and enter them into the

statistician's data shell. Neither teachers nor students were contacted to gain information related to this study.

Data collection instrument. The instrument used to collect data was a data shell in an Excel spreadsheet. Each row in the file included individual student data for his or her fourth-grade SOL reading score and fifth-grade SOL reading score. Once data were gathered, a research assistant entered scores with anonymous identifiers. Data were identified by "AR" and "Non-AR."

Instrument measures. The measurement scores for this study were derived from the Virginia SOL reading test. The fourth-grade results were used as the pretest and the fifth-grade scores were used as the posttest. Reliability and validity of the SOL tests are shown in the *Virginia Standards of Learning Technical Report* (2010), which was completed following the 2008–2009 administration of the tests. The accepted reliability coefficient was .70 using Cronbach's alpha. For the fifth-grade reading test, the reliability coefficient was .82 for paper administration and .84 for online administration. With the exception of a small number of special education students, all students took the online version of the reading test in school year 2011–2012.

Content validity provides evidence that the tests are appropriate measures of the content taught. This validity is proven through examination of the intended instructional material (grade-level SOLs) along with the curriculum blueprint provided to each educator that instructs them on the specific standards to appear on the end-of-year assessment as well as the number of questions for that standard. In addition to the blueprint, teachers receive a pacing guide for each content area and a curriculum framework to be used to design their instruction. These tools provide support for the

validity of the SOL tests by providing appropriate instructional material to ensure that the content taught is also tested.

Data analysis procedures. Initially, it was expected that the pretest scores would differ significantly which would point to ANCOVA as the method of analysis. ANCOVA would have allowed the comparison of posttest means while taking into account initial differences. According to Ary et al. (2006), this adjustment improves the precision of the experiment and, "with the part of the variance in the posttest scores that is *not* [italics added] caused by treatment removed, any difference caused by the treatment stands out more clearly" (p. 308). However, when using an independent samples *t* test, it was found that pretest scores were not significantly different; in fact the pretest scores were nearly identical.

This finding indicated that the appropriate statistical method for analysis of the posttest scores was that of independent *t* tests. Borg, Gall, and Gall (1993) state that the *t* test is used to establish if the observed difference between the mean of scores of "two groups on a measure is likely to have occurred by chance or whether it reflects a true difference in the mean scores of the populations represented by the two groups" (p. 158).

Data, in the form of end-of-year SOL scores, did not identify specific students at any school site. Once the test data were entered into a spreadsheet by a research assistant, names were deleted and students' scores were listed as either "AR" or "Non-AR" and used anonymous identifiers. The analysis of the test scores was to determine if there was a statistically significant difference in the performance on the SOL reading test for students who consistently used the AR program compared to those who did not use the program.

CHAPTER FOUR: FINDINGS

Chapter Four includes a discussion of the purpose of the study, the method chosen for data analysis, the research questions and hypotheses, as well as the results. The dependent variable in this study was students' SOL reading scores for the fifth-grade 2012 administration of the tests. The covariate for this study was the 2011 SOL reading test scores for the same students, which also functioned as the pretest. The independent variable was the mandatory use of AR. Data were analyzed to determine, first, if there was a statistically significant difference in the overall performance of students in either treatment group and, second, if there was a statistically significant difference when the data were examined based on gender in both treatment groups. Results were considered significant at $p = .05$. When the value of Sig. > .05, it indicates that the corresponding statistic may have occurred by chance, thus increasing the possibility of a Type I error.

Purpose of the Study

The purpose of this causal-comparative study was to add to the body of literature analyzing the possible effectiveness of AR on fifth-grade students' reading comprehension as measured by their performance on the SOL reading test. The use of AR in classrooms can require spending a significant amount of instructional funds. This researcher felt that it was important for administrators and teachers to have access to as much evidence as possible to determine the program's level of effectiveness before investing those scarce funds.

Study Overview

This study took place in a school system in Virginia whose total student population is approximately 14,000 students, preschool through 12th grade. The

population studied was that of a random sampling of fifth-grade students from two instructional treatments. One group of students was instructed using only the school system's literacy plan, while the other group was instructed using the literacy plan along with the mandatory use of AR. The study used archived SOL reading test data from 2011 (fourth grade) and 2012 (fifth grade).

Data Analysis and Methodology

Data from this study were analyzed using independent *t* tests. The independent *t* test is useful in causal-comparative, or ex post facto, research because it assists in determining if the difference in the mean scores of two groups happened by chance or if the independent variable (AR) could have caused the difference (Borg et al., 1993)

Data were gathered by the researcher through access to archived end-of-year test scores for 2012. Once the data for each student were entered into a spreadsheet, a research assistant entered the scores into a data shell that provided anonymous identifiers for each score. Data were analyzed using independent *t* tests. The means for each group were determined and analyzed both by group (reading method) and by group/gender.

Results

Research Question 1: What is the difference between the 2012 Virginia Standards of Learning reading scores in fifth-grade students who used the system literacy plan along with mandatory use of Accelerated Reader as compared to the scores of fifth-grade students who used only the literacy plan?

The results of this study, as shown through the use of independent *t* tests, indicated that there was no statistically significant difference between the SOL reading scores of students who were instructed using only the school system-required literacy

plan and students who were instructed using the school system-required literacy plan along with mandated use of AR. Due to this result, the researcher failed to reject the null hypothesis.

Analysis by reading method. Table 2 shows the descriptive statistics for both groups on the 2011 pretest and the 2012 posttest.

Table 2

Descriptive Statistics by Group

Reading method		M	SD	n
SOL reading 2011	AR	490.31	74.920	51
	Non-AR	498.86	62.772	57
	Total	494.82	68.582	108
SOL reading 2012	AR	490.29	62.282	51
	Non-AR	487.51	61.047	57
	Total	488.82	61.360	108

Table 2 contains the means of student scores from both the 2011 and 2012 administration of the SOL reading test. Analysis of the 2011 administration showed that the AR group mean was 490.31 and the non-AR group mean was 498.86 indicating that the non-AR group performance was higher than the AR group at pretest. The mean for both groups for the 2012 administration showed that neither group significantly increased their reading performance—the AR group's mean score changed from 490.31 to 490.29 and the non-AR group from 498.86 to 487.51, with the AR group dipping only slightly.

When analyzing the data by reading method, it was determined that neither group showed statistically significant growth. In actuality, when examined by group, neither showed growth by any meaningful measure. The group that was mandated to use AR in addition to the literacy plan showed a decrease in the mean of .02 points between the

2011 SOL reading scores and the 2012 SOL reading scores. However, the non-AR group showed a decrease in the mean of 11.35 points during the same period. Although seemingly significant when simply looking at the decrease in the mean, it does not allow for a determination of any effect attributed to the use of AR based on statistical significance. The independent *t* test for SOL reading by group (see Table 3) indicated that neither reading method showed statistically significant results over time with $p=.05$. Based on these results, the decision was made to fail to reject the null hypothesis for Research Question 1.

Table 3

Independent t tests for by Method 2012

Group	n	M	SD	t	p
AR	51	490.29000	62.28200		
				0.2340	0.8154
Non-AR	57	487.51000	61.04700		

Research Question 2: What is the difference between the 2012 Virginia Standards of Learning reading scores by gender for fifth-grade students who used the system literacy plan along with mandatory use of Accelerated Reader as compared to the scores of fifth-grade students who used only the literacy plan?

When examining the data using independent *t* test to determine whether gender affected the results, it was determined that there was, again, no statistically significant difference when comparing girls and boys in both treatment groups. Due to this result, the decision was made to fail to reject both null hypotheses related to differences in performance by gender.

Analysis by gender. Independent *t* tests were done to determine if there were any differences by gender for those students who used AR plus the Literacy Plan during their fifth grade year as compared to those students who used the Literacy Plan alone. Table 4 shows the results of these independent *t* tests between females in each group.

Table 4

Independent t tests by Gender(Female)

Group	n	M	SD	t	p
AR	26	483.04000	65.58200		
				0.8299	0.4107
Non-AR	24	498.00000	61.542.00		

Table 4 shows that when analyzing using independent *t* tests, the reading SOL score means for the female AR group and the female Non-AR group did not show a statistical difference at the end of fifth grade with significance at p=.05. Based on these results, the decision was made to fail to reject null hypothesis 2.

Table 5

Independent t tests by Gender (Male)

Group	n	M	SD	t	p
AR	25	497.84000	55.36800		
				1.1610	0.2505
Non-AR	33	479.88000	60.47400		

Table 4 shows that when analyzing using independent *t* tests, the reading SOL score means for the male AR group and the male Non-AR group did not show a statistical

difference at the end of fifth grade with significance at p=.05. Again, based on these results, the decision was made to fail to reject null hypothesis 3.

The results of the examination of the data as a group by gender brought results that surprised this researcher. At the elementary level, it is often assumed that girls are more apt to read than boys and girls are more apt to be successful with reading than boys (Below, Skinner, Fearrington, & Sorrell, 2010; Logan & Johnston, 2009; Logan & Johnston, 2010; McIntosh, Reinke, Kelm, & Sadler, 2013; Prado & Plourde, 2011). Table 6 indicates that although girls began with higher SOL reading scores in 2011 compared to boys, their scores dropped, while male scores increased. One possibility for this change in scores is that the scores for girls exhibited regression to the mean with the 2012 administration of the SOL reading test. The 2011 mean, at 507.26, was quite high, with that score falling within the Pass/Advanced range (500–600).

Table 6

Descriptive Statistics by Gender

Source	Gender	*M*	*SD*	*n*
SOL reading 2011	Female	507.26	68.178	50
	Male	484.10	67.683	58
	Total	494.82	68.582	108
SOL reading 2012	Female	490.22	65.073	50
	Male	487.62	58.518	58
	Total	488.82	61.360	108

When examining the data by treatment, the results again were surprising. The mean for the boys who participated in AR increased 16.32 points over the time period,

while the mean for the girls who participated in AR decreased by 15.73 points. In the non-AR group, the mean for boys decreased by 6.18 points, while the mean for girls decreased 18.46 points. Again, this change in the mean is not considered statistically significant yet from an instructional perspective, the results require further examination. Further discussion will be covered in Chapter Five.

Summary

The purpose of this study was to answer two questions: What is the difference between the 2012 Virginia Standards of Learning reading scores in fifth-grade students who used the system literacy plan along with mandatory use of Accelerated Reader as compared to the scores of fifth-grade students who used only the literacy plan? and What is the difference between the 2012 Virginia Standards of Learning reading scores by gender for fifth-grade students who used the system literacy plan along with mandatory use of Accelerated Reader as compared to the scores of fifth-grade students who used only the literacy plan?

The results of this study indicated that there was no statistical verification that the use of AR had any influence on the reading comprehension scores of fifth-grade students when examined by group or by gender. Therefore, the findings indicate that there was sufficient evidence to fail to reject the null hypotheses for both research questions. The significance of these results and the instructional implications of the data are discussed in further detail in Chapter Five.

CHAPTER FIVE: DISCUSSION

The ESEA of 1965 was the beginning of federal control in the local classroom. This far-reaching act was established along with other social legislation in conjunction with President Lyndon Johnson's Great Society, which was his response to the growing number of Americans in poverty (Hewitt, 2011). The purpose of the legislation was to provide federal funding to local school districts to help offset the cost of educating large populations of students in poverty. The ESEA began the practice of tying federal funding to schools through various programs such as Title I, which provided financial assistance for remediation for low-income students (Hewitt, 2011).

The ESEA was revised during the presidency of George W. Bush and renamed the No Child Left Behind Act of 2001. The legislation requires schools to ensure that all students are reading on grade level by 2014 (NCLB, 2002). The accountability requirements placed on every teacher because of the changes in NCLB (2002) have sent them searching for that magic bullet that will help students become better readers, without impacting the time teachers need to teach the other content areas in the school day (Blankstein, 2011; Hansen et al., 2009; D. W. Johnson & Johnson, 1999; Nunnery & Ross, 2007; Oppenheimer, 2007; Pavonetti et al., 2003; Taylor et al., 2005). AR is one such program whose website (Renaissance Learning, 2012) makes promises to students, teachers, and parents that have yet to be proved.

This study was conducted in response to this researcher's experience while examining the research provided by Renaissance Learning (2012) on its website to support the success of its flagship product, Accelerated Reader. Prior to this research, this researcher's measure of the effectiveness of AR had been measured through

anecdotal information gained from the teachers within the six schools in two states in which she had worked. Although never a supporter of the mandated use of AR, this researcher faced a situation in her own school when the fifth-grade teachers intended to mandate the use of AR and make its use a portion of each student's grade. This requirement for the mandated use of AR posed an ethical dilemma for this researcher/administrator. She believes, as other researchers have confirmed, that children will become strong readers when the instruction of reading intrigues and engages them. Reading success and satisfaction does not develop when students' reading lists are based on AR book availability rather than student interest (Esposito & Smith, 2006; Guthrie, et al., 2004; McKool, 2007).

The initial review of the literature began with a visit to the Renaissance Learning website where, after entering the appropriate search terms, the results provided 166 studies. The review of the studies indicated that of the 166 studies listed, 65 were merely anecdotal reports from teachers or administrators; 44 were articles and summaries produced or funded by Renaissance Learning; and 23 were conducted by researchers not affiliated with Renaissance Learning, with the exception of two written by Topping, who is affiliated with Paul (creator of AR). The remaining 34 studies were reviewed (along with 284 other studies related to AR) by the WWC. Only two of the 318 total studies considered by the WWC met its review criteria, and neither study accepted showed statistically significant effects on student reading fluency or comprehension.

An additional concern related to the literature reviewed was that nearly all of the studies were more than 7 years old, with many written in the late 1980s and early 1990s. This lack of significant effects in any study reviewed, paired with the lack of current data,

encouraged this researcher to design a study that would attempt to provide current data in a treatment group that utilized AR in such a way that its results can be compared to a control group with similar demographics.

Summary of Findings

Research design. This study took place in a Virginia school system with approximately 14,000 students. The study used the SOL reading test scores of two groups of fifth-grade students. The design of this study is that of causal-comparative, or ex post facto, which works well when examining students in preexisting groups. This allows the researcher to look back on behavior or performance and attempt to determine the influence of a certain variable (Schenker & Rumrill, 2004). Since it was not possible to divide the students into a control and treatment group prior to the study, student scores were randomly chosen from intact groups. Student scores from schools that used only the school system's literacy plan were used as the control group, while student scores from schools that used the literacy plan along with the mandated use of AR were used as the treatment group. Scores were gathered by the researcher from archived data for the end of school year 2011 and the end of school year 2012. Individual scores were then entered into a data shell by a research assistant, at which time anonymous identifiers replaced student names. Data were then analyzed using independent *t* tests, which were run twice: once to determine any difference in reading methods between groups and again to determine any difference in reading methods based on gender. When using p=.05 as the level of significance, independent *t* tests run comparing reading method by group resulted in p=0.8154. When examining the data by gender and group, using p=.05, the independent *t* test for females was p=0.4107 and for males was p=0.2505. None of

the independent *t* tests found statistical significance when comparing the use of Accelerated Reader along with the use of the system Literacy Plan and the use of the Literacy Plan alone.

Discussion of findings. The score range for Virginia SOLs is as follows: Failing = <400, Pass/Proficient = 400–499, and Pass/Advanced = 500–600. The descriptive statistics for SOL reading by reading method and group indicated that the fourth-grade reading scores for both groups were quite close and quite high, with the AR group mean at 490.31 and the non-AR group mean at 498.86. Both group means were just below the Pass/Advanced range.

The descriptive statistics also showed that the mean score for both groups decreased over time. The 2012 group mean for the AR group decreased from 490.31 to 490.29, while the group mean for the non-AR group decreased from 498.86 to 487.51. Since the group means were initially high, this decrease could be the result of regression to the mean, but the decrease begs further analysis from an instructional standpoint. The statistical lens through which this study was examined indicated that the use of the AR reading method when studied by group or by gender did not have a positive effect on student reading comprehension.

In comparison to past research on AR, this study's results do not support the findings of studies that showed statistical increases in student reading comprehension (e.g., R. A. Johnson & Howard, 2003; Magnolia Consulting, 2010; Melton et al., 2004; Pappas et al., 2010; Vollands et al., 1999). It should be noted that even though Melton et al.'s 2004 study showed that AR had a moderate effect on student comprehension, when the program ended, so did student reading. The discrepancy between the results of this

study and past studies advocates for future researchers to continue to examine groups of students in a control group and experimental group setting. Future studies will help to determine if this study's results are applicable only to this setting or if, in fact, the use of AR has no effect on student comprehension.

However, these results may support the studies that indicated that participation in and success with the AR program is seen more at lower grade levels. Several studies (Nunnery & Ross, 2007; Nunnery et al., 2006; Topping & Paul, 1999; Topping et al., 2007a) indicated that as students entered upper grades, their success and participation in AR lessened. It seems that when students reach upper elementary and middle school, their desire to participate in AR diminishes. This lessening of the desire to read was also seen in a study conducted by Stockard (2010) related to direct instruction for students to prevent the "fourth-grade slump" (p. 148). This slump at such a pivotal time in reading can have a profound effect on student comprehension of fifth-grade reading material. The fact that, overall, the means of both groups decreased over time causes this researcher to ponder the questions that arose during the literature review: Do students rebel at being forced to read a contrived list of books that have a mandatory multiple-choice test attached (Esposito & Smith, 2006; Luck, 2010; G. Thompson et al., 2008)? Do students miss the true value of reading when they have little control over what they read (Biggers, 2001; Cuddeback & Ceprano, 2002; McKool, 2007; G. Thompson et al., 2008)?

AR process. The original intent of this study was to determine, in a specific setting, if the use of the AR program had any effect on fifth-grade students' reading comprehension as measured by their performance on the end-of-year SOL reading test.

The AR process begins when a student takes a reading placement test using the STAR computerized test that determines his or her reading level, called the ZPD. The concept of this level is loosely linked to Vygotsky's ZPD, which originally defined the process of a teacher and student working closely together to learn a skill or concept, with the teacher slowly lessening his or her influence and the student becoming independent with the activity.

Once a student's ZPD is determined, teachers give students a grade-level range from which to choose books. Each book on the AR list has a grade-level equivalency as well as a point value. When the student completes a book from the AR list, he or she will take a 5- to 20-question quiz that will provide immediate feedback on the percentage of questions answered correctly and the number of points accrued. All schools in which this researcher has worked followed this process and kept a record of student performance on AR. At the end of the reporting period, students were rewarded in some tangible way (bookmarks, coupons, pizza parties, etc.) depending on the number of points accrued. Students who did not meet their AR goal were either reprimanded for their failure to meet the goal or ignored because of their less-than-stellar efforts.

It is this process of recognition that segregates high-performing readers from low-performing readers that caused this researcher to question the benefits of, first, purchasing only those books that have accompanying AR tests (at $3 per test) and, second, steering students' book choices to only those books on the AR list. Another nagging concern comes from those who questioned the benefits of AR and the long-term advantages of bribing students through the use of tangible incentives to read (Facemire, 2000; Retelsdorf et al., 2011; G. Thompson et al., 2008).

Theoretical framework and AR. AR uses its interpretation of Vygotsky's ZPD as the theoretical framework for the program. This researcher's study of the original intent of the ZPD revealed that Vygotsky indicated that, in the relationship between the teacher and the learner, the teacher's influence is constant throughout the learning process. The teacher is seen as a partner in learning who gradually moves away to allow for independence as the student gains mastery. The teacher maintains a close relationship through continued monitoring to ensure that the mastery continues as the student gains ownership over the skill or concept (Vygotsky, 1978). The AR version of this relationship-strong theory involves having the student's reading competency level measured by a computer-generated test that provides the teacher with a grade-level equivalency range. This range is used to set reading goals for students for a defined time period. The teacher provides each student with his or her ZPD range and then monitors, by computer, their progress. AR's version of the ZPD seems to require teachers to spend more time with computers than with students.

The spirit of Vygotsky's (1978) ZPD pairs well with effective reading instruction. Researchers have found that effective reading instruction involves teachers possessing a strong knowledge of their students' abilities to not only read text but also be able to comprehend, connect, and use what they have read. In order to have this knowledge of student abilities, reading must be taught in a social context that involves deep discussions about what has been read (Allington, 2005; Andreassen & Braten, 2011; Brozo, 2010; Esposito & Smith, 2006; Glasswell & Ford, 2011; Guthrie et al., 2004). In a study conducted by Pavonetti et al. (2003), the researchers found, in talking with teachers, that students were discouraged from discussing books they had read for fear that other

students would learn enough about the book to take and pass an AR test without actually having read the book.

Instructional Implications

The statistical results of this study showed no statistically significant effect when comparing the use of only the literacy plan to the use of the literacy plan along with mandated use of AR. Although not statistically significant, this researcher believes that the results of the study require that the school system review the data to determine why the mean for the students who did not use AR dropped 11.35 points from the 2011 administration of the SOL reading test to the 2012 administration of the test. Even more dramatic was the dip in the mean for female students. The mean for female non-AR students' test scores dropped 18.46 points (from 516.46 to 498.00). The mean for female AR students dropped 15.73 points (from 498.77 to 483.04). Male non-AR students' mean dropped 6.18 points (from 486.06 to 479.88), while male AR students' mean showed the only increase, with a rise of 16.32 points. The argument for regression to the mean could certainly be used to explain the decrease in the mean for SOLs by group and for girls. The means were very high at the end of fourth grade (2011), yet it is still an instructional concern that growth was not seen by group or by gender (except for boys).

Fullan, Hill, and Crevola (2006) described the old mission of schools as that of providing access for all students to a basic education that prepared students for the workplace or, for a small group of students, access to a university education. The authors described the new mission, however, as having been foisted upon the public education system and that of a continuation of the old mission, with a shift to giving all students access to an education that is lifelong and is able to adapt to whatever careers the

technologies of this new century will present. In order to provide this type of education for students in classrooms today, teachers must find what motivates students internally to want to continue learning long after grades have been given. Researchers have provided insight into this often-ignored area of teacher training and have found that intrinsic motivation is a strong predictor of student performance (Cameron & Pierce, 1994; Dai & Wang, 2007; Deci & Ryan, 2008; Dixon-Krauss, 1995; Guthrie et al., 2007; Husman et al., 2005; Law, 2008; Logan et al., 2011; Park, 2011; Putman & Walker, 2010).

Facemire (2000) found that when teachers ignore the importance of intrinsic motivation and depend on extrinsic motivators, they devalue the importance of becoming a good reader. A 2007 study by Guthrie et al. determined that teachers should recognize that students' interest in what they read is a personal investment. Therefore, teachers should work with students to help them hone their interests and improve their ability to choose books that will increase their reading performance rather than depending on students' desire for a reward.

Any positive effects from extrinsic motivation are short-lived. Melton et al. (2004) found that although reading increased with the use of prizes and rewards, when the motivators stopped being given, so did the reading. Both Facemire (2000) and Houston (2008) believed that when students were given contrived reading lists based on a reading range rather than interest, they failed to develop any lifelong strategies for selecting appropriate reading material. And, when examining how students felt about the reading requirements of AR, Esposito and Smith (2006) found that students did not like being forced to read and that if the rewards were removed, they would no longer read. As for low-performing readers, Logan et al. (2011) found in their study that motivation

explained more of the variance among the poor readers than did their reading skill. The effects of an extrinsically motivated reading program can also have long-lasting effects. A study of middle school students found that students who read the most came from elementary schools that did not use AR (Pavonetti et al., 2003).

In this researcher's study, the means of SOL reading scores for both AR and non-AR groups were high with the 2011 administration of the test. At the 2012 administration of the test, the means for both groups dropped, with the non-AR group dropping the most. Although the high scores indicated that the system literacy plan is helping students be successful with reading (means for both groups were at or above the Pass/Advanced range), the fact that the scores dropped indicated that students need continued instruction and encouragement to not only maintain their current level of proficiency, but to show gains each year.

In order for students to become strong, independent readers, teachers must change their reading instruction. Reading instruction should include research-based strategies that have been shown to increase the amount of students' reading, as well as increased their comprehension and their desire to read. Teachers need to make certain that they are providing students with sufficient prior knowledge before reading, whether as a class or individually (Alexander & Jetton, 2000; Anmarkrud & Braten, 2009; Brozo, 2010; Glasswell & Ford, 2011; Guthrie et al., 2007; Hilden & Pressley, 2007; Katz & Assor, 2006; Pecjak, Podlesak, & Pirc, 2011; Retelsdorf et al., 2011; Samuelstuen & Braten, 2005; Taboada et al., 2009). Prior knowledge involves discussions with students as a class before texts are read in order to provide a foundation upon which new knowledge can be added. It also involves individual discussions with students to determine interests

and background before making reading list recommendations. When students begin a new text with even a cursory knowledge of the subject matter, they are better able to make connections with self, connections with text, and connections with the world. These connections are vital if students are to make sense of what they read (Keene & Zimmermann, 1997).

Teachers also need to recognize that high-ability readers generally read the most and the best (Borman & Dowling, 2004; FCRR, 2006; Krashen, 2005). When competition is created to encourage reading, students who are already considered good readers will flourish, while struggling readers will still struggle—a concept known as the Matthew Effect. Reading performance will improve when teachers begin to provide reading experiences that will interest and engage students, while also providing clear, explicit instruction in reading strategies (Dixon-Krauss, 1995; Ferguson et al., 2011; Hilden & Pressley, 2007; Law, 2008; Prado & Plourde, 2011; Stockard, 2010). It is after these strategies are taught to students that the intended use of the ZPD (Vygotsky, 1978) is of the most benefit. Once strategies are taught to students, it is the teacher's role to monitor their use of strategies until students truly own them. Once a teacher finds that students are consistently using the reading strategies that have been taught, then he or she can retreat and trust that his or her influence and instruction have produced good readers.

Conclusion

The results of this study indicated that although the use of AR showed no statistically significant benefit for students, its use did no harm. However, the program is quite expensive, especially with the company's move to a completely online format, which could cost this researcher's school several thousand dollars per year to maintain.

If research does not support a high return on investment for this product, why should schools and school systems continue to fund it? A better use for those funds may be to hire extra hands for the classroom in order to allow the teacher more time with smaller groups of students. It is during this time, with a well-trained professional, that students will genuinely learn to love reading and experience the joy of success.

Limitations

This study is limited based on several factors. The study reviewed data from random sampling of intact groups as opposed to groups that were randomly assigned to each treatment. The dependent variable may limit generalizing the results to schools located in Virginia due to the use of the Virginia SOL reading test. Another limitation is the fact that the researcher was employed at one of the schools in the study. However, the benefit of conducting a causal-comparative study is the lack of interference on the part of the researcher. Teachers managed their reading classes without the knowledge that their students' performance would be examined. Also, students had no knowledge that their scores would be used, thus eliminating any possibility of the Hawthorne Effect (Chiesa & Hobbs, 2008).

In order to prevent any interference with the fifth-grade classes in the school at which this researcher worked, teachers were not made aware of the details of the study until the testing was completed.

Summary and Recommendations for Future Research

Hundreds of presumed studies have been conducted related to the effectiveness of AR. The issue with the studies is that most were not conducted under experimental or even quasiexperimental conditions, which makes their results suspect. Many of the

researchers noted that the lack of fidelity of implementation between teachers may have had an impact on study results. It is believed that this study resolved some of those concerns. Future research should continue to compare a group of students who consistently use AR to a control group of students who use a consistent, uniform curriculum in order to isolate which reading method has the most positive effect on students' reading comprehension. An additional recommendation would be to compare the effect of Accelerated Reader on a lower grade group's progress with that of an upper grade group's progress. This comparison would attempt to determine if AR does have more effect on the reading comprehension of younger students as opposed to that of older students.

After reviewing the results of this study, this researcher regrets not adding a qualitative aspect that provided information from students related to their feelings about AR. The results of this study indicated that only male students who used AR showed any growth over time. The analysis of the results could have been deeper had there been some indication of why female students' performance dropped so sharply over time, while non-AR male students' performance remained virtually flat and AR male students' performance rose slightly. Was this the result of the pretest scores being artificially high at the end of fourth grade or was it regression to the mean? It is believed that interviewing the students could have answered those questions.

In order for this researcher's school to convert to the online-only version of AR, the budget would need to include a yearly cost of nearly $2,500. For this researcher, the results from this study do not support the continued investment in purchasing only those books for which there is an AR test. The results also indicate that fifth-grade teachers as

well as administrators need to examine the current reading program and determine what changes need to be made to the instructional program to increase reading interest for male students. The reading habits of female students also need to be examined to determine why their scores dropped so dramatically over time.

REFERENCES

Abrami, P., Cholmsky, P., & Gordon, R. (2001). *Statistical analysis for the social sciences: An interactive approach.* Needham Heights, MA: Allyn & Bacon.

Alexander, P. A., & Jetton, T. L. (2000). Learning from text: A multidimensional and developmental perspective. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 285–310). Mahwah, NJ: Erlbaum.

Alexander, P. A., & Murphy, P. K. (1998). Profiling the difference in students' knowledge, interest, and strategic processing. *Journal of Educational Psychology, 90,* 435–447.

Allington, R. L. (2005, February). Ideology is still trumping evidence. *Phi Delta Kappan,* 462–468.

Alvermann, D. (2003). *Seeing themselves as capable and engaged readers: Adolescents and re/mediated instruction.* Naperville, IL: Learning Point.

Andreassen, R., & Braten, I. (2011). Implementation and effects of explicit reading comprehension instruction in fifth-grade classrooms. *Learning and Instruction, 21,* 520–537.

Anmarkrud, O., & Braten, I. (2009). Motivation for reading comprehension. *Learning and Individual Differences, 19,* 252–256. doi:10.1016/j.lindif.2008.09.002

Ary, J., Jacobs, L., Razavieh, A., & Sorensen, C. (2006). *Introduction to research in education* (7th ed.). Belmont, CA: Wadsworth.

Below, J., Skinner, C., Fearrington, J., & Sorrell, C. (2010). Gender differences in early literacy: Fifth-grade Dynamic Indicators of Basic Early Literacy Skills probes. *School Psychology Review, 39*(2), 240-257.

Benner, G. J., Nelson, J. R., Ralston, N. C., & Mooney, P. (2010). A meta-analysis of the effects of reading instruction on the reading skills of students with or at risk of behavioral disorders. *Behavioral Disorders, 35*(2), 86–102.

Biggers, D. (2001). The argument against Accelerated Reader. *Journal of Adolescent & Adult Literacy, 45*(1), 72–75.

Blankstein, A. M. (2010). *Failure is not an option: Six principles for making student success the only option* (2nd ed.). Thousand Oaks, CA: Corwin.

Blankstein, A. M. (2011). *The answer is in the room: How effective schools scale up student success.* Thousand Oaks, CA: Corwin.

Borg, W., Gall, J., & Gall, M. (1993). *Applying educational research: A practical guide* (3rd ed.). White Plains, NY: Longman.

Borman, G. D., & Dowling, N. M. (2004). *Testing the Reading Renaissance program theory: A multilevel analysis of student and classroom effects on reading achievement.* Unpublished manuscript, University of Wisconsin-Madison. Retrieved from http://www.education.eisc.edu/elpa/people/faculty/Borman /BormanDowling2004.pdf

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Brem, S. K., Husman, J., & Duggan, M. A. (2005). Findings from a three-year study of Reading Renaissance in a Title I urban elementary school. *Academic Exchange Quarterly, 9*(3), 355–359.

Brozo, W. G. (2010). The role of content literacy in an effective RTI program. *Reading Teacher, 64*(2), 147–150. doi:10.1598/RT.64.2.11

Bullock, J. (2005). *Effects of the Accelerated Reader on reading performance on third, fourth, and fifth grade students in one western Oregon elementary school* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3181085)

Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research, 64*(3), 363–383.

Carnine, D. (2000). *Why education experts resist effective practices (and what it would take to make education more like medicine).* Retrieved from http://www .edexcellence.net/doc.carnine.pdf

Carter, B. (1996). Hold the applause! Do Accelerated Reader and Electronic send the right message? *School Library Journal, 42*(2), 22–25.

Chiesa, M., & Hobbs, S. (2008). Making sense of social research: How useful is the Hawthorne Effect? *European Journal of Social Psychology, 38,* 67–74. doi:10 .1002/ejsp.401

Clark, C., & Osborne, S. (2007). Demystifying the reluctant reader. *Literacy Today, 53,* 22–23.

Cuddeback, M. J., & Ceprano, M. A. (2002). The use of Accelerated Reader with emergent readers. *Reading Improvement, 39*(2), 89–94.

Dai, D. Y., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology, 32*, 332–347. doi:10.1016/j.cedpsych.2006.05.002

Daniel, S., Walsh, A, Goldston, D., Arnold, E., Reboussin, B., & Wood, F. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities, 39*(6), 507–514.

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology, 49*(3), 182–185. doi: 10.1037/a0012801

Denton, C., & Al Otaiba, S. (2011). Teaching word identification to students with reading difficulties and disabilities. *Focus on Exceptional Children, 43*(7), 1–15.

Dixon-Krauss, L. A. (1995). Partner reading and writing: Peer social dialogue and the zone of proximal development. *Journal of Reading Behavior, 2*(1), 45–63. doi:10 .1080/10862969509547868

Dzaldov, B. S., & Peterson, S. (2005). Book leveling and readers. *Reading Teacher, 59*(3), 222–229. doi:10.1598.RT.59.3.2

Esposito, J., & Smith, S. (2006). From reluctant teacher to empowered teacher-researcher: One educator's journey toward action research. *Teacher Education Quarterly, 33*(3), 45–57.

Facemire, N. E. (2000). *The effect of the Accelerated Reader on the reading comprehension of third graders* (Master's thesis). Available from ERIC database. (ED442097)

Ferguson, N., Currie, L., Paul, M., & Topping, K. (2011). The longitudinal impact of a comprehensive literacy intervention. *Educational Research, 53*(3), 237–256. doi: 10.1080/00131881.2011.598657

Florida Center for Reading Research. (2006). *Accelerated Reader.* Retrieved from www .fcrr.org/fcrrreports/pdf/acceleratedreader.pdf

Fullan, M., Hill, P., & Crevola, C. (2006). *Breakthrough.* Thousand Oaks, CA: Sage.

Gall, M., Borg, W., & Gall, J. (1996). *Educational research: An introduction* (6th ed.). White Plains, NY: Longman.

Glasswell, K., & Ford, M. (2011). Let's start leveling about leveling. *Language Arts, 88*(3), 208–216.

Goodman, G. (1999). *The Reading Renaissance/Accelerated Reader program* (Pinal County School-to-Work evaluation report). Available from ERIC database. (ED427299)

Groce, R., & Groce, E. (2005). Deconstructing the Accelerated Reader program. *Reading Horizons, 46*(1), 18–30.

Guthrie, J. T., Hoa, A. L. W., Wigfield, A., Tonks, S. M., Humenick, N. M., & Littles, E. (2007). Reading motivation and reading comprehension growth in the later elementary years. *Contemporary Educational Psychology, 32,* 282–313. doi:10 .1016/j.cedpsych.2006.05.004

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K., Taboada, A., Davis, M., . . . Tonks, S. (2004). Increasing reading comprehension and engagement through Concept-Oriented Reading Instruction. *Journal of Educational Psychology, 96*(3), 403–423. doi:10.1037/0022-0663.96.3.404

Hansen, L. E., Collins, P., & Warschauer, M. (2009). Reading management programs: A review of the research. *Journal of Literacy and Technology*, *10*(3), 56–80.

Harper, C. (2001). *A study of Accelerated Reader model and master schools and performance on the Mississippi Curriculum Reference Test* (Unpublished master's thesis). Retrieved from http://www.geocities.ws/gautierlibrary/ARstudy.htm

Heinemann. (2010). *Fountas and Pinnell A to Z Benchmark Assessment System.* Retrieved from http://heinemann.com/fountasandpinnell/research /BASFieldStudyFullReport.pdf

Hewitt, D. (2011). Reauthorize, revise, and remember: Refocusing the No Child Left Behind Act to fulfill *Brown*'s promise. *Yale Law & Policy Review, 30*(1), 169–194.

Hilden, K. R., & Pressley, M. (2007). Self-regulation through transactional strategies instruction. *Reading & Writing Quarterly, 23*, 51–75. doi:10.1080 /10573560600837651

Houston, C.R. (2008). The use of reading levels as alternative classification in school libraries. *Cataloging & Classification Quarterly, 45*(4), 65–80. doi:10.1300 /J104v45n04_05

Husman, J., Brem, S., & Duggan, M. A. (2005). Student goal orientation and formative assessment. *Academic Exchange Quarterly*, *9*(3). Retrieved from http://www .thefreelibrary.com/Student+goal+orientation+and+formative+assessment. -a0138703719

Investors' Business Daily. (2011). [Home page.] Retrieved from http://www.investors .com

Jenkins, J., & Terjeson, K. J. (2011). Monitoring reading growth: Goal setting, measurement frequency, and methods of evaluation. *Learning Disabilities Research & Practice, 26*(1), 28–35.

Johnson, D. W., & Johnson, R. (1999). *Learning together and alone: Cooperative, competitive, and individualistic learning* (5th ed.). Boston, MA: Allyn & Bacon.

Johnson, R. A., & Howard, C. A. (2003). The effects of the Accelerated Reader program on the reading comprehension of pupils in grades three, four, and five. *Reading Matrix, 3*(3), 87–96.

Kambarian, V. N. (2001). *The role of reading instruction and the effect of a reading management system on at-risk students* (Unpublished doctoral dissertation). Available from ERIC database. (ED461835)

Katz, I., & Assor, A. (2006). When choice motivates and when it does not. *Educational Psychology Review, 19*(4), 429–442. doi:10.1007/s10648-006-9027-y

Keene, E., & Zimmermann, S. (1997). *Mosaic of thought: Teaching comprehension in a reader's workshop.* Portsmouth, NH: Heinemann.

Knapp, M. (2008). How can organizational and sociocultural learning theories shed light on district instructional reform? *American Journal of Education, 114*, 521–539.

Kohn, A. (2004). NCLB and the effort to privatize public education. In D. Meier & G. Wood (Eds.), *Many children left behind: How the No Child Left Behind Act is damaging our children and our schools* (pp. 79–97). Boston, MA: Beacon Press.

Kozulin, A. (2002). Sociocultural theory and the mediated learning experience. *School Psychology International, 23*(1), 7–35.

Krashen, S. (2005). Accelerated Reader: Evidence still lacking. *Knowledge Quest, 33*(3), 48–49.

Kravtsova, E. (2009). The cultural-historical foundations of the zone of proximal development. *Journal of Russian and East European Psychology, 47*(6), 9–24.

Lane, K. L., Little, M. A., Redding-Rhodes, J., Phillips, A., & Welsh, M. T. (2007). Outcomes of a teacher-led reading intervention for elementary students at risk for behavioral disorders. *Exceptional Children, 74*(1), 47–70.

Law, Y. K. (2008). The role of attribution beliefs, motivation and strategy use in Chinese fifth-graders' reading comprehension. *Educational Research, 51*(1), 77–95. doi: 10.1080/00131880802704764 2

Logan, S. & Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where these differences lie. *Journal of Research in Reading, 32*(2), 199-214. doi: 10.1111/j.1467-9817.2008.01389.x

Logan, S. & Johnston, R. (2010). Investigating gender differences in reading. *Educational Review, 62*(2), 175-187. doi: 10.1080/00131911003637006

Logan, S., Medford, E., & Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers reading comprehension performance. *Learning and Individual Differences, 21*(1), 124–128.

Luck, S. A. (2010). Accelerated Reader: The controversy continues—A literature review of the effectiveness of Accelerated Reader in increasing reading achievement and student motivation. *Illinois Reading Council Journal, 39*(2), 3–9.

Macalister, J. (2011). Today's teaching, tomorrow's text: Exploring the teaching of reading. *ELT Journal, 65*(2), 161–169. doi:10.1093/elt/ccq023

Macaruso, P., & Walker, A. (2008). The efficacy of computer-assisted instruction for advancing literacy skills in kindergarten children. *Reading Psychology, 29,* 266–287. doi:10.1080/02702710801982019

Magnolia Consulting. (2010). *A final report for the evaluation of Renaissance Learning's Accelerated Reader program.* Charlottesville, VA: Author.

McCardle, P., & Chhabra, V. (2004). *The voice of evidence in reading research.* Bethesda, MD: Brookes.

McIntosh, K., Reinke, W., Kelm, J., & Sadler, C. (2013). Gender differences in reading skill and problem behaviors in elementary school. *Journal of Positive Behavior Interventions, 15*(1), 51-60. doi: 10.1177/1098300712459080

McKnight, D. (1992). *Using the Accelerated Reader and other strategies and varied techniques to improve the reading attitudes of fifth grade students.* Available from ERIC database. (ED350582)

McKool, S. (2007). Factors that influence the decision to read: An investigation of fifth grade students' out-of-school reading habits. *Reading Improvement, 44*(3), 111–113.

Meier, D. (2004). NCLB and democracy. In D. Meier & G. Wood (Eds.), *Many children left behind: How the No Child Left Behind Act is damaging our children and our schools* (pp. 66–78). Boston, MA: Beacon Press.

Melton, C. M., Smothers, B. C., Anderson, E., Fulton, R., Replogle, W. H., & Thomas, L. (2004). A study of the effects of the Accelerated Reader program on fifth grade students' reading achievement growth. *Reading Improvement, 41*(1), 18–23.

Milone, M. (2008). *The development of ATOS: The Renaissance readability formula.* Wisconsin Rapids, WI: Renaissance.

Monroe, L. (2005). *Evening the score: Accelerated Reader and aesthetic stance* (Master's thesis). Retrieved from http://students.ou.edu/M/L.A.Monroe-1/artifact _EveningScore.pdf

Moore, D., Bean, T., Birdyshaw, D., & Rycik, J. (1999). *Adolescent literacy: A position statement for the Commission on Adolescent Literacy of the International Reading Association.* Newark, DE: International Reading Association.

Neale, M. (1989). *Neale analysis of reading ability* (Rev. British ed.). Windsor, England: NFER-Nelson.

Niewolny, K., & Wilson, A. (2009). What happened to the promise? *Adult Education Quarterly, 60*(1), 26–45.

No Child Left Behind Act of 2001, Pub. L. No. 107-110 (2002).

Nunnery, J. A., & Ross, S. M. (2007). The effects of school Renaissance program on student achievement in reading and mathematics. *Research in the Schools, 14*(1), 40–59.

Nunnery, J. A., Ross, S. M., & McDonald, A. (2006). A randomized experimental evaluation of the impact of Accelerated Reader/Reading Renaissance implementation on reading achievement in grades 3 to 6. *Journal of Education for Students Placed At Risk, 11*(1), 1–18. Retrieved from http://www.memphis.edu /crep/pdfs/Accelerated _Reader_JESPAR_11__1-18.pdf

Oppenheimer, T. (2007). Selling software: How vendors manipulate research and cheat students. *Education Next, 7*(2), 22–30.

Palmiter, A. S., Arcaira, E. R., White, R. N., & Reisner, E. R. (2009). *The literacy programs of Save the Children: Results from the 2008–09 school year.* Washington, DC: Policy Studies Associates.

Pappas, D. N., Skinner, C. H., & Skinner, A. L. (2010). Supplementing Accelerated Reader with classwide interdependent group-oriented contingencies. *Psychology in the Schools, 47*(9), 887–902. doi:10.1002/pits.20512

Park, Y. (2011). How motivational constructs interact to predict elementary students' reading performance: Examples from attitudes and self-concept in reading. *Learning & Individual Differences, 21,* 347–358.

Patall, E. A., Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. *Psychological Bulletin, 134*(2), 270–300. doi:10.1037/0033-2909.134.2.270

Paul, T. D. (1996). *Patterns of reading practice.* Wisconsin Rapids, WI: Renaissance.

Pavonetti, L. M., Brimmer, K. M, & Cipielewski, J. F. (2003). Accelerated Reader: What are the lasting effects on the reading habits of middle school students exposed to Accelerated Reader in elementary grades? *Journal of Adolescent & Adult Literacy, 46*(4), 300–311.

Pecjak, S., Podlesak, A., & Pirc, T. (2011). Model of reading comprehension for 5th grade students. *Studia Psychologica, 53*(1), 53–67.

Pitcher, S., Martinez, G., Dicembre, E., Fewster, D., & McCormick, M. (2010). The literacy needs of adolescents in their own words. *Journal of Adolescent & Adult Literacy, 53*(8), 636–645. doi:10.1598/JAAL.53.8.2

Prado, L., & Plourde, L. (2011). Increasing reading comprehension through the explicit teaching of reading strategies: Is there a difference among genders? *Reading Improvement, 48*(1), 32–43.

Putman, M., & Walker, C. (2010). Motivating children to read and write: Using informal learning environments as contexts for literacy instruction. *Journal of Research in Childhood Education, 24*(2), 140–153.

Renaissance Learning. (2007). *Getting results with Accelerated Reader.* Wisconsin Rapids, WI: Author.

Renaissance Learning. (2010). *Getting the most out of STAR reading: Using data to inform instruction and intervention.* Wisconsin Rapids, WI: Author.

Renaissance Learning. (2012). *Accelerated Reader.* Retrieved from www.renlearn.com/ar/default.aspx

Retelsdorf, J., Koller, O., & Moller, J. (2011). On the effects of motivation on reading performance growth in secondary school. *Learning and Instruction, 21,* 550–559. doi:10.1016/j.learninstruc.2010.11.001

Rodriguez, S. (2007). The Accelerated Reader program's relationship to student achievement on the English-language arts California Standards Test. *Reading Matrix, 7*(3), 191–205.

Ross, S. M., Nunnery, J., & Goldfeder, E. (2004). *A randomized experiment on the effects of Accelerated Reader/Reading Renaissance in an urban school district: Final evaluation report.* Retrieved from http://www.research.renlearn.com/research/pdfs/322.pdf

Rudd, P., & Wade, P. (2006). *Evaluation of Renaissance Learning mathematics and reading programs in UK specialist and feeder schools.* Retrieved from http://www.nfer.ac.uk/nfer/publications/SRY01.pdf

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25,* 54–67. doi:10.1006/ceps.1999.1020

Sadusky, L. A., & Brem, S. K. (2002). *The integration of Renaissance programs into an urban Title I elementary school, and its effect on school-wide improvement* (2001–2002 Annual Report to Renaissance Learning). Retrieved from http://www.research.renlearn.com/research/139.asp

Samuels, S. J., Lewis, M., Wu, Y. C., Reininger, J., & Murphy, A. (2003). *Accelerated Reader vs. non-Accelerated Reader: How students using the Accelerated Reader outperformed the control condition in a tightly controlled experimental study.* Retrieved from http://www.tc.umn.edu/~samuel001/web%20pdf/Final%20Report--Accelerated%20Reader%20vs%20Non-Accelerated%20Reader.pdf

Samuels, S. J., & Wu, Y. C. (2003). *The effects of immediate feedback on reading achievement* (Technical Report). Minneapolis: University of Minnesota.

Samuels, S. J., & Wu, Y. C. (2004, May). *How the amount of time spent on independent reading affects reading achievement.* Retrieved from http://research.renlearn.com/research/189.asp

Samuelstuen, M., & Braten, I. (2005). Decoding, knowledge, and strategies in comprehension of expository text. *Scandinavian Journal of Psychology, 46,* 107–117.

Schacter, J. (2000). *Reading programs that work: A review of programs for pre-kindergarten to 4th grade* (Report funded by the Milken Family Foundation). Retrieved from http://www.mff.org/pubs/ME279.pdf

Schenker, J., & Rumrill, P. (2004). Causal-comparative research designs. *Journal of Vocational Rehabilitation, 21,* 117–121.

Schmidt, R. (2008). Really reading: What does Accelerated Reader teach adults and children? *Language Arts, 85*(3), 202–211.

Schraw, G., & Bruning, R. (1999). How implicit models of reading affect motivation to read and reading engagement. *Scientific Studies of Reading, 3*(3), 281–302.

Smith, A., & Westberg, K. (2011). Student attitudes toward Accelerated Reader: "Thanks for asking!" *Current Issues in Education, 14*(2), 1–7. Available from ERIC database. (EJ938968)

Souto-Manning, M. (2010). Accelerating reading: Inequities in the early years. *Language Arts, 88*(2), 104–113. Available from ERIC database. (EJ906326)

Stanfield, G. (2006). *Incentives: The effects on reading attitude and reading behaviors of third-grade students.* Unpublished master's thesis, Georgia College and State University, Milledgeville.

Stanovich, K. E. (1986). Matthew Effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360–407.

Stockard, J. (2010). Promoting reading achievement and countering the "fourth grade slump." The impact of direct instruction on reading achievement in fifth grade. *Journal of Education for Students Placed at Risk, 15,* 218–240.

Taboada, A., Tonks, S. M., Wigfield, A., & Guthrie, J. T. (2009). Effects of motivational and cognitive variables on reading comprehension. *Reading and Writing: An Interdisciplinary Journal, 22*(1), 85–106.

Taylor, B., Pearson, P., Peterson, D., & Rodriguez, M. (2005). The CIERA school change framework: An evidence-based approach to professional development and school reading improvement. *Reading Research Quarterly, 40*(1), 40–69.

Thompson, G., Madhuri, M., & Taylor, D. (2008). How the Accelerated Reader program can become counterproductive for high school students. *Journal of Adolescent & Adult Literacy, 51*(7), 550–560. doi:10.1598/JAAL.51.7.3

Thompson, M. (2012). *Honors English.* Retrieved from http://www.davis.k12.us/ffjh /thompson/arcon.pdf

Topping, K. J. (2006). *Accelerated Reader in specialist schools.* Retrieved from http:// research.renlearn.com/research/257.asp

Topping, K. J., & Paul, T. (1999). Computer-assisted assessment of practice at reading: A large scale survey using Accelerated Reader data. *Reading & Writing Quarterly, 15*, 213–231.

Topping, K. J., Samuels, J., & Paul, T. (2007a). Computerized assessment of independent reading: Effects of implementation quality on achievement gain. *School Effectiveness and School Improvement, 18*(2), 191–208. doi:10.1080 /09243450601104834

Topping, K. J., Samuels, J., & Paul, T. (2007b). Does practice make perfect? Independent reading quantity, quality and student achievement. *Learning and Instruction, 17*, 253–264. doi:10.1016/j.learninstruc.2007.02.002

Topping, K. J., Samuels, J., & Paul, T. (2008). Independent reading: The relationship of challenge, non-fiction and gender to achievement. *British Educational Research Journal, 34*(4), 505–524. doi:10.1080/01411920701609380

Topping, K. J., & Sanders, W. L. (2000). Teacher effectiveness and computer assessment of reading: Relating value added and learning information system data. *School Effectiveness and School Improvement, 11*(3), 305–337.

University of Dundee. (2012). [Home page.] Retrieved from http://www.dundee.ac.uk /eswce/people/kjtopping.htm

U.S. Department of Education. (2012). *Definitions.* Retrieved from http://eddataexpress .ed.gov/definitions.cfm

Virginia Department of Education. (2010). *Virginia Standards of Learning technical report, 2010.* Retrieved from http://www.doe.virginia.gov/testing/test _administration/technical_reports/sol_technical_report_2008-09_administration _cycle.pdf

Virginia Department of Education. (2012a). *Accreditation and AYP reports.* Retrieved from http://www.doe.virginia.gov/statistics_reports/accreditation_ayp_reports _ayp/index.shtml

Virginia Department of Education. (2012b). *Testing and Standards of Learning.* Retrieved from http://www.doe.va.gov/testing/sol/standards_docs/english/review .shtml

Vollands, S. R., Topping, K. J., & Evans, R. M. (1999). Computerized self-assessment of reading comprehension with the Accelerated Reader: Action research. *Reading and Writing Quarterly, 15,* 197–211.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Vygotsky, L. (1986). *Thought and language.* Cambridge, MA: MIT Press.

Walberg, H. J. (2001). *Final evaluation of the reading initiative: Report to the J. A. & Kathryn Albertson Foundation board of directors.* Retrieved from http://research .renlearn.com/research/130.asp

Walter, D. G. (2010). [Home page.] Retrieved from http://damiengwalter.com

Wang, J. H., & Guthrie, J. T. (2004). Modeling the effects of intrinsic motivation, extrinsic motivation, amount of reading, and past reading achievement on text comprehension between U.S. and Chinese students. *Reading Research Quarterly, 39*(2), 162–186.

What Works Clearinghouse. (2010). [Home page.] Retrieved from www.renlearn.com/ar /default.aspx

White, R. N., & Reisner, E. R. (2007). *Model literacy programs: Save the Children. Evaluation findings from the 2005–06 school year.* Available from ERIC database. (ED498796)

Wolters, C. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist, 38*(4), 189–205.

APPENDIX A: INSTITUTIONAL REVIEW BOARD EXEMPTION

LIBERTY
UNIVERSITY.

The Graduate School at Liberty University

October 17, 2012

Jan S. Nichols
IRB Exemption 1428.101712: Accelerated Reader and Its Effects on Fifth-Grade Students' Reading Comprehension

Dear Jan,

The Liberty University Institutional Review Board has reviewed your application in accordance with the Office for Human Research Protections (OHRP) and Food and Drug Administration (FDA) regulations and finds your study to be exempt from further IRB review. This means you may begin your research with the data safeguarding methods mentioned in your approved application, and that no further IRB oversight is required.

Your study falls under exemption category 46.101 (b)(4), which identifies specific situations in which human participants research is exempt from the policy set forth in 45 CFR 46:

(4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

Please note that this exemption only applies to your current research application, and that any changes to your protocol must be reported to the Liberty IRB for verification of continued exemption status. You may report these changes by submitting a change in protocol form or a new application to the IRB and referencing the above IRB Exemption number.

If you have any questions about this exemption, or need assistance in determining whether possible changes to your protocol would change your exemption status, please email us at irb@liberty.edu.

Sincerely,

Fernando Garzon, Psy.D.
Professor, IRB Chair
Counseling

(434) 592-4054

LIBERTY
UNIVERSITY.