

Investigating the Effectiveness of a Reading Recovery Program  
for At-Risk Students in an Alternative School Setting

by  
Clyde E. Alderman, Jr.

An Applied Dissertation Submitted to the  
Fischler School of Education and Human Services  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Education

Nova Southeastern University  
2008

Approval Page

This applied dissertation was submitted by Clyde E. Alderman, Jr. under the direction of the persons listed below. It was submitted to the Fischler School of Education and Human Services and approved in partial fulfillment of the requirements for the degree of Doctor of Education at Nova Southeastern University.

---

Paul Borthwick, PhD  
Committee Chair

---

Date

---

Robert Kershaw, EdD  
Committee Member

---

Date

---

Maryellen Maher, PhD  
Executive Dean for Research and Evaluation

---

Date

## Abstract

Investigating the Effectiveness of a Reading Recovery Program for At-Risk Students in an Alternative School Setting. Alderman, Clyde E., Jr., 2008: Applied Dissertation, Nova Southeastern University, Fischler School of Education and Human Services. High Risk Students/Child Behavior/Reading Research/Grade 7/Grade 8

This applied dissertation was designed to evaluate a Reading Recovery program at an alternative school in Florida. The Reading Recovery program was used as an intervention for improving students' reading skills and performances on the Florida Comprehensive Assessment Test (FCAT) that is given each year to students in Grades 7 and 8.

The researcher used the students' scores on the FCAT reading and English language arts tests to evaluate the effectiveness of the Reading Recovery program as an intervention in the researchers' school district. Each year, students were placed in the remedial reading class because they failed to achieve a passing score on the test in the previous years. Prior to this study, no data had been available to evaluate the effectiveness of a Reading Recovery program of at-risk students' academic performances.

As a result of this applied dissertation, the researcher sought to assist the school in becoming more effective and efficient in implementing the Reading Recovery program. This study represents the school's attempt to evaluate the effectiveness of a Reading Recovery program implemented by every teacher in Grades 7 and 8. The evaluation results suggested that a Reading Recovery program was successful in reducing the number of at-risk students. The researcher was able to determine that the school had decided what needed to be done in order to improve the at-risk students' FCAT scores. Also, the researcher found that the majority of the teachers implemented the program as it was designed. The findings also suggested that the teachers who participated in the program viewed it as beneficial. The findings further indicated that the teachers now view assessments as an important piece in identifying individual student needs.

## Table of Contents

	Page
Chapter 1: Introduction .....	1
Statement of the Problem.....	1
Significance of the Study .....	3
Purpose of the Study .....	4
Definition of Terms.....	4
Limitations .....	5
Chapter 2: Literature Review .....	7
Theoretical Models of Literacy Acquisition .....	7
Brief History of Reading Instruction .....	8
Interventions Pertinent to This Applied Dissertation.....	10
Profile of Chapter 1 Model .....	11
Profile of the New Zealand Reading Recovery Model .....	12
Brief History of Reading Recovery in the United States.....	13
Impact of Reading Recovery on the Research Community .....	14
Individual Instruction.....	15
Instructional Emphasis.....	16
Staff Development .....	19
Time Allocation .....	20
Student Selection .....	21
Program Effectiveness and Student Performance.....	22
Summary .....	23
Chapter 3: Methodology .....	25
Research Design.....	25
Populations.....	26
Designation of Groups .....	26
Reading Instruction.....	27
Selection Process .....	27
Program Description .....	28
Reading Recovery .....	28
Chapter I.....	29
Measurement Instruments .....	31
Norms.....	32
Reliability.....	32
Validity .....	33
Classroom Reading Inventory.....	33
The Stanford 9 Achievement Test .....	34
Data Collection .....	36
Data Analysis and Procedures .....	37
Chapter 4: Data Analysis .....	38
Profile of the Populations.....	38
Attrition.....	38

Total Number of Student Participants.....	39
Study Participants .....	40
Results.....	40
Summary of the Test Populations.....	42
Characteristics of Reading Recovery Students.....	43
Characteristics of Reading Recovery and Chapter I Students .....	45
Characteristics of Chapter I Students.....	49
Characteristics of Wait-Listed Students.....	50
 Chapter 5: Discussion, Implications, Recommendations, and Conclusion .....	 51
Discussion.....	53
Implications for Practice.....	56
Recommendations for Further Research.....	58
Conclusion .....	60
 References .....	 61
 Tables	
1 Student Participation in Supplemental Reading Programs in the 2006-2007 School Year.....	38
2 Attrition of Program Participants in the 2006-2007 School Year.....	39
3 Student Eligibility for Inclusion in Study in 2006.....	40
4 Student Participation in Study in 2007 .....	41
5 Summary of Florida Comprehensive Assessment Test Scores: Three Populations .....	42
6 Summary of Florida Comprehensive Assessment Test Scores: Reading Recovery Populations .....	44
7 Details of Florida Comprehensive Assessment Test Scores: Number of Years in Reading Recovery .....	45
8 Summary of Florida Comprehensive Assessment Test Scores: Comparison Populations .....	46
9 Details of Florida Comprehensive Assessment Test Scores: Number of Years in Chapter I .....	48

## Chapter 1: Introduction

### *Statement of the Problem*

Reading is a problem-solving message getting activity that swells in power and versatility the more that it is rehearsed (Clay, 1998). One of the major tasks put before educators today is making sure that every student learns how to read fluently and independently and comprehends what is being read. Still, under the most perfect conditions, there are students who struggle with learning to read irrespective of the teacher, the strategy or the method employed.

Middle school students across the nation arrive at school with assorted and wide-ranging, primary learning experiences and capabilities that, combined, can make reading a difficult undertaking. The range of student abilities and needs within the classroom require an array of instructional strategies and methods. Over time, researchers and educators have tried to ascertain the efficacy of various reading strategies and methods in helping students learn to become proficient and independent readers (Slavin, 1995). Those attempts have provided the foundation for discussions and debates on whether corrective or intervention programs have positive or negative effects on reading outcomes for students. Among the discussions and debates, there is an extensive concurrence among researchers that early intervention is preferred to remediation (Aldridge, 2000).

In response to the foregoing problem, the No Child Left Behind (NCLB) Act of 2001 was introduced by President George Bush to evaluate literacy skills among students (U.S Department of Education, 2005). The NCLB Act of 2001 required annual testing of all public school students in reading and mathematics in Grades 3 to 8 and high school by the 2005-2006 school year; annual report cards on school performance for parents, voters, and taxpayers ensuring that each child reads by the third grade; and highly qualified

teachers being in every public school classroom by 2006. Under the NCLB Act of 2001 Reading First initiative, funds are provided to states and school districts to eliminate reading deficits in kindergarten to Grade 3 by establishing and implementing high-quality, scientifically research-based reading instructional programs (U.S. Department of Education). Teachers and school officials also have flexibility and control in choosing how to use federal education funds. Reading First is a national initiative of the NCLB Act of 2001 that is designed to make every child in the United States a proficient reader by the 2013-2014 school year. In response, states and school district personnel are also required to increase the overall percentage of students who are reading at or above grade level. To fulfill these requirements, all children must be tested annually in reading, and the results of these assessments will provide clear evidence of Reading First's effectiveness (U.S. Department of Education). This research project addressed the problem of low literacy skills and the impact of the research-based Reading Recovery program that was implemented at a lower income, alternative middle school to improve the reading levels of the students.

According to the reading test scores and data collected from the school's reading department, 60% of the Grade 7 students performed below grade level on the reading section of the Florida Comprehensive Assessment Test (FCAT). The FCAT was designed to measure student achievement of Florida's Sunshine State Standards (Florida Department of Education, 1996). The results of the tests were reported in terms of FCAT mean scores and FCAT reading achievement levels.

It is against this background that the administrators and reading teachers of this middle school implemented a Reading Recovery program with a view to improve reading achievement among seventh-grade students during the spring school year. The

intervention was an intensive, instructional reading program that combined research-based instructional strategies with quality professional development for teachers. This Reading Recovery program has been available for use in classrooms across the country since 1999. Research showed that this Reading Recovery program has had positive effects on students' reading performances. The problem was that no information existed regarding the effectiveness of the Reading Recovery program. The administrators of the program implemented the Reading Recovery program to evaluate and improve at-risk students' reading scores.

### *Significance of the Study*

The underlying premise of the Reading Recovery program was that early intervention in middle school is critical for students at risk of reading failure. According to the Reading Recovery Council of North America (2001), "This is supported by longitudinal research that shows that children who fail behind grade one tend to remain below grade level in later school years" (p. 1). Slavin (2005) observed that, whereas success in the early grades does not guarantee success throughout the school years and beyond, failure in the early grades does virtually guarantee failure in later schooling.

Among the emergent literacy and intervention programs studied such as Reading Recovery and Success for All have been shown to be highly successful with below level elementary school students (Allington, Stuetzel, Shake, & Lamarche, 2000; Clay, 1998; Donmoyer & Kos, 1996; Perfetti & Curtis, 1995; Pinnell, 1996; Reutzel, 1999). However, comprehensive research on effective reading instruction programs for adolescents who are at risk is lacking at this present time.

This study contributes to the adolescent literacy knowledge base. According to VanKleeck (1998), the literacy development of a 12-year-old in middle school or a 17-

year-old in high school is as serious a concern to society as the literacy development of a preschool child or a child in the elementary grades. Thus, without serious changes in planning and implementation of curriculum programs, not much will change to support and extend the literacy development in these older students.

### *Purpose of the Study*

The purpose of this study was to identify criteria for and to evaluate the effectiveness of the Reading Recovery program for at-risk, middle school students. Based on the results of the evaluation and the literature review, the results were utilized for program improvement and will serve as a future guide for teachers and administrators. The researcher addressed one research question: What will be the effectiveness of a Reading Recovery program on improving the FCAT scores of at-risk students?

Much effort on the part of central office administrators, reading coordinators, principals, reading specialists, teachers, and a parent liaison went into the establishment of the Chapter 1 program. This applied dissertation also examined data that could be helpful in planning its future curricula or in program evaluation.

### *Definition of Terms*

The following terms are defined for this study.

*Reading Recovery program.* This is an early intervention procedure designed to accelerate the reading progress of the lowest achieving ninth graders and to prevent future failure.

*Syntax.* This is a study of the pattern of formation of sentences and phrases from words and of the rules for the formation of grammatical sentences in a language.

*Stanford 9.* This is the particular edition of the Stanford Achievement Test administered to seventh-, eighth-, and ninth-grade students in the school district.

*Title and Chapter I.* These are interchangeable terms that describe the federal government's largest initiative to narrow the achievement gap between low-income and high-income students.

*Wait Listed.* This refers to those students in the bottom half of the class who are eligible but not in need of instruction because of relatively high scores.

### *Limitations*

The findings of this applied dissertation were site specific. Therefore, the focus of its test questions may not directly match with universal Reading Recovery program goals and objectives.

Student participants were preselected on the basis of their need for reading remediation through a combination of teacher assessment and formal and informal testing. Thus, participation in the program was not based on random assignment to a study group or a comparison group.

The populations included all seventh graders who were eligible for supplemental reading instruction that had Stanford 9 scores below grade level. No standardized test results were available for the students under study in Grades 6 or 7 due to school district testing philosophy that exempts those grade levels from testing, so students in these grades were excluded. Student attrition also affected population size; for example, Stanford 9 test results were unavailable for students who had moved out of the county by spring or who had chosen home or private schooling. Some students who had been held back one grade level also had no Stanford 9 scores, and some children who had begun the Reading Recovery program were later identified as special education students and were not included in the study. Other students had Stanford 9 scores that were either unavailable for analysis or were not administered under standard conditions. As a result,

a total of 400 of the 600 students remained in alternative school, seventh-grade classrooms at the end of the 9 weeks' grading period.

## Chapter 2: Literature Review

This chapter summarizes theories related to the acquisition of literacy, traces the history of reading instruction, and profiles Chapter I. It describes the New Zealand reading recovery model and its American importation and explores the impact of the reading recovery model on the research community.

### *Theoretical Models of Literacy Acquisition*

Chard, Kame'enui, and Simmons (1994) agreed that the goal of reading instruction is to produce independent readers with “clear communication, strong reading comprehension, articulates writing, and critical analysis skills” (p. 142). However, the optimal settings and techniques that best accomplish this goal are controversial. A wide range of researchers, including linguists, sociologists, anthropologists, economists, humanists, and traditional neuroscientists, investigated a number of specialized components of the reading process. Their diverse perspectives, added to those of the cognitive and educational psychologists, simultaneously enrich and confuse issues. Invariably, their findings can be categorized into two philosophical models, reductivism and constructivism, that sharply divide the professional reading community.

In the first model, reading is reduced to parts that respond to, according to Spiegel (as cited in Deegan, 1995), “explicit instruction, guided practice, and independent application” (p. 692). The systematic relationship that exists between letters and sounds gives readers a frame of reference for identifying unknown words. Sometimes labeled the *proficient model* (Singer & Ruddell, 1997), it has the approval of practitioners and researchers who support a code emphasis. It is held that development proceeds in a linear progression from letter acquisition to decoding of words to comprehension of word pronunciation.

In the second model, according to Giroux (as cited in Deegan, 1995), reading is considered a process that flows from comprehension “to critique to emancipation [of the self]” (p. 689). A reader’s previous knowledge and critical interaction with text produce comprehension prior to or in lieu of perceptions of individual words. Sometimes labeled the *development model*, constructing and reshaping personal meaning continuously in top-down mode is considered, according to Goodman (1998), “part of a [somewhat] discontinuous progressive movement in education ranging back through the language experience movement to the ideas of Dewey and other 20th-century progressives” (p. 1539).

A third model, referred to 20 years ago as *interactive* (Rumelhart, 1996) and more recently as *eclectic* (Stahl, McKenna, & Pagnucco, 1995) or *balanced* (Fitzgerald, 1999), rejects rigid adherence to either previously described model. The developmental model blends visual and phonemic elements into a strong decoding program in the early grades while stressing comprehension strategies and literature discussion to promote understanding of text and love of reading as fluency increases. The interactive model combines explicit phonics instruction with an authentic learning task and literature and provides a focus on achievement to incorporate the most salient components of both models. These models are important for students to improve their reading, pronunciation, and comprehension levels. Without the methodology of models, students lose interest in reading.

### *Brief History of Reading Instruction*

Because reading is central to academics and key to maximal learning, public policy regarding reading has long demonstrated the capacity to stir emotions and create tension among teachers, researchers, and the public. National trends in reading instruction

seem to mirror current applications of educational psychology to the general learning process in somewhat cyclical fashion. The advent of the scientific movement in education in the late 1930s introduced a rational system to the studies of instructional technique and classroom management that filtered into reading practice.

The 1950s, late 1970s to 1980s, and the late 1990s to early 2000s can be characterized as periods in which there was support for holistic practices, integrated curriculum, open classroom for collaborative growth, and whole language in education and personal learning. In the early 1970s, basic process research focused on language acquisition. By the mid-1980s, reading had become inseparably linked with its companion language components--writing, listening, and speaking and also with metacognitive processes. As Goodman (1998) urged, readers should become “not . . . perceivers of text but as . . . questioners of text” (p. 1537). With research grounded in humanism, the whole language model valued reading behaviors that typified an ideal egalitarian society (McKenna, Stahl, & Reinking, 1995) and instilled the need for assessments to evaluate individual growth through test scores.

By contrast, the postwar 1940s to early 1960s and early 1980s marked periods of direct instruction, planned curriculum, and a view of students as passive receptors of learning. Because, as stated by MacLean (1997), “the pronunciation of English is related in a regular way to its writing system” (p. 516), phonics became a mainstay of these reform periods. Chall (1996), a leading spokeswoman for the letter-sound approach, claimed that research demonstrated that “earlier, more systematic phonics produced better word recognition and reading comprehension . . . through third grade . . . than did later less systematic instruction” (p. 1539). The product approach evident in the 1950s basal focus on literal comprehension continued with the large-scale studies of beginning

reading in the 1960s and again into the early 1980s with the pursuit of text-related variables such as vocabulary to test reading achievement.

Elements of both instructional philosophies continued to shape reading practices throughout the last decade of the century. Emergent literacy that is deemed a continuum, rather than a distinct stage of readiness-to-read, encompasses a broad spectrum of experiences with printed text as well as with specific areas of knowledge. Emergent literacy embraces a period of development in which the informal interactions of very young children with oral and written language establish the foundation for their subsequent reading and writing acquisition.

Phonemic awareness, an auditory skill that traditionally has been considered a correlate of phonics, is foremost among emergent literacy behaviors. Recent research by VanKleeck (1998) reiterated the importance of direct instruction in the letter-sound relationship to expedite word recognition. The inclusion of phonics in child-centered instruction is currently spreading, and authentic assessment practices continue to increase.

#### *Interventions Pertinent to This Applied Dissertation*

Despite best practice, a subset of children who struggle with language activities has inevitably been apparent as early as kindergarten. Although these students frequently received differentiated instruction in kindergarten as well as support from various school personnel, they often failed to attain proficiency in reading in subsequent years. Through careful, long-term observation, formal and informal assessment by teachers and specialists and conferences with parents, schools ushered these children into remediation, retention, or special education on a case-by-case basis. Thus, students improve on their reading and comprehension skills through intervention.

*Profile of Chapter I Model*

Of the three alternatives available to low-achieving students, remediation through the federal initiative known as Chapter I is the oldest and most widespread option. As an outgrowth of President Lyndon B. Johnson's Great Society, Congress passed the Elementary and Secondary Education Act in 1965 to channel additional resources to impoverished children. The largest program was Chapter I or Title I that, according to Allington and Walmsley (1995), "provided funds for the education of students who were considered hard to teach and in need of additional educational resources that few districts routinely provided" (p. 25).

An in-class setting required the reading specialist to function as an aide to students in completion of class work or to introduce new or extra resources to teachers. A pullout setting allowed the reading specialist to focus on small group instruction (Allington et al., 2000; Bean, Eichelberger, Lazar, & Zigmond, 1995; Carter, 1996; Juel, 1996; Slavin, 1995; Thistlethwaite & Mason, 1995). Positive results were small at best and, often, nonexistent (Allington & Walmsley, 1995). Reauthorization of Chapter I in 1988 and 1994 attempted to mandate increased accountability and classroom structure. Over the last decade, several programs attempted to combine the most powerful features of remedial reading instruction to accelerate the academic progress of at-risk children. Both phonemic awareness studies as early as 1983 and comprehension strategy studies of the mid-1980s to early 1990s demonstrated that low-achieving, young students could progress satisfactorily with concrete instruction and practice (Pressley, 1995).

Although some investigators focused on singular, familiar dimensions of literacy acquisition, others sought to optimize progress through more ambitious reforms. Two innovative programs, introduced in the 1980s, were the most notable: Reading Recovery

developed by Clay of New Zealand and Success for All, developed by Slavin of Johns Hopkins University in Baltimore, Maryland. Each initiative featured an instructional blend of individual tutoring by trained teachers, elements of phonics, and whole language and also included training for educational staff. However, the differences between the two were striking. Success for All was a comprehensive, preventive effort that regrouped entire grade levels of students by ability during the common reading block, provided support for families in need, and implemented system-wide changes in language arts instruction and assessment. Reading Recovery, an early intervention, focused intensely on first graders only and offered intensive, pullout instruction for students, training for teachers, and extended opportunities for collegial support.

Each innovation had supporters as well as detractors, but Reading Recovery alone established a network for rapid dissemination of techniques that continues to this day. Marketed aggressively in Virginia at reading and language conferences and by word of mouth, Reading Recovery, as implemented by the school district, was the focus of this applied dissertation. Teachers are astonished at how Reading Recovery has helped their at-risks students to become better readers and improve on their comprehension skills.

Reading Recovery has promoted parental involvement with their children's reading progress.

#### *Profile of the New Zealand Reading Recovery Model*

In 1962, developmental psychologist, Clay, applied methods for studying and recording child behavior to observations of children during their 1st year of school in Auckland, New Zealand. At primary schools chosen from high, middle, and lower income areas, children were randomly selected from class rolls in order to examine the relationships between instruction, text, and language acquisition. The study concluded in

1966 with new questions and a need for further research.

By the mid-1970s, Clay (1996a) pilot tested the program in five schools with 122 children and “ordinary teachers released for individual teaching” (p. 173) to conduct a follow-up study. Students not tutored formed the control groups and were taught by the same teachers using the same lessons. Initial, final, and follow-up testing data were collected on book level and reading vocabulary variables. Clay (1996a) concluded that discontinued students had maintained their gains 1 and 2 years later, whereas those students not discontinued predictably required further tutoring. Clay (1996a) further inferred that the control and discontinued students had absorbed the program’s processing strategies defined loosely as those automatic habits leading to reading independence.

#### *Brief History of Reading Recovery in the United States*

Researchers at Ohio State University observed the New Zealand program onsite in 1983. With funding from two foundations and the Ohio State University, in 1984-1985, the Columbus Public Schools and the Ohio Department of Education jointly imported Clay and the national director of Reading Recovery in New Zealand, Watson, to train Reading Recovery teachers.

A pilot group of six Columbus Public Schools preceded the programs first full year of operation. For research purposes, 136 children and 32 trained Reading Recovery teachers, joined by an alternative, compensatory reading program group and a random sample of first graders, were grouped together in September 1985.

Tested on text reading ability at the end of second, third, and fourth grades, the discontinued Reading Recovery group displayed higher mean scores than the comparison group. They scored within the average band of the random sample for 2 years and at the bottom of the average band in the 3rd year. Ohio State University researchers were

satisfied that Reading Recovery was an effective intervention for short-term and long-term. Shortly thereafter, Reading Recovery expanded to include 33 schools in 10 Ohio school districts with a total of 400 first graders participating in one of five different intervention programs for purposes of comparison.

After 14 weeks, test results of four reading instruments (Text Reading Level, Dictation Assessment Task, Woodcock Reading Mastery, and Gates-McGinitie) demonstrated that Reading Recovery children performed better on all tests than the children in other prevention programs. Retested at the end of Grade 1 and again at the beginning of Grade 2, Reading Recovery students alone sustained such performances. Researchers concluded that those discontinued students were well within the average band, had maintained their ability to profit from mainstream classroom instruction, and had continued to learn. Ohio State University became synonymous with Reading Recovery, anchoring the North American Reading Recovery program and copyrighting the name. By 1987, North American educators from outside Ohio began to train at Ohio State University, and phenomenal growth ensued. By 1991-92, 38 states in the United States and four Canadian provinces were participating in Reading Recovery programs. Within a few years, 49 states had operative programs, 9 states offered Reading Recovery in Spanish at 50 sites, and more than 460,000 children had been served since the program's inception (Askew, Fountas, Lyons, Pinnell, & Schmitt, 1998).

#### *Impact of Reading Recovery on the Research Community*

Research into the Reading Recovery program tended to focus largely on the effectiveness of the initiative itself or that of specific design features. The earliest Reading Recovery articles to emerge from studies of the initial American implementation of the program in Ohio riveted the attention of the large national reading community.

Pinnell (1996) reported that pilot study data and 1st full year data in Columbus, Ohio, provided evidence of Reading Recovery's short-term and long-term success with low-achieving first graders. Conducted in part by colleagues Knight and Fischer (1995), the study revealed that Reading Recovery students performed better than a control group and compared favorably with a random sample group on seven of nine Clay diagnostic measures. Monitored the 2nd year, former Reading Recovery students had maintained their skills in text reading.

Pinnell's (1996) study burst upon practitioners and researchers ripe for direction. Because much of the available literature had provided only updates of the whole language versus phonics issue (Stahl et al., 1995) or consisted of qualitative analyses of observational studies, the quantitative results of the New Zealand and Ohio reading recovery models immediately piqued the interest of researchers. Studies such as Stahl et al.'s search to compare and contrast Reading Recovery with extant intervention programs for youngsters or those replicating Pinnell's study at Reading Recovery sites were now expanding explosively.

Researchers also scrutinized the allocations of time, student selection, and instructional content. Each feature of Reading Recovery was analyzed, acclaimed, and refuted at least once throughout the decade. The program's instructional emphasis on the research design itself produced the most comments.

### *Individual Instruction*

One-on-one teaching is the signature Reading Recovery vehicle for expressing Clay's (1996b) strong commitment to the goal of student independence in reading and writing. It enables a tutor to reinforce positive student responses and refute and redirect negative thought processes immediately. By tailoring Reading Recovery lessons to the

unique strengths and weakness of students, tutors extend all learning from the context of what each particular child understands about reading at any given point in time. Cognition of Pinnell's (1996) work that showed that individualized instruction in Reading Recovery exceeded results associated with small-group instruction led Juel (1996) to search for elements that contributed to the power of tutoring. The study indicated that three components mattered: a supportive relationship between tutor and student; significant scaffold experience in letter-sound instruction; and continuous, clear modeling of decoding and encoding processes. Ross (1995) contrasted the progress of students tutored within Reading Recovery with those tutored with individual instruction. Ross reported significant, overall effect sizes on passage in recovery's tutored students but nonsignificant results on word identification and word attack.

Hobsbaum (1996) focused on the effect of tutoring on writing behaviors in Reading Recovery. Juel's (1996) study supported Clay's (1996a) premise that daily tutoring provides the comprehensive view of a student, which maximizes success, but the results suggested that its suitability may best serve short-term, task-specific goals such as accruing reading strategies, rather than sustained comprehension and writing processes.

### *Instructional Emphasis*

The instructional emphases of Reading Recovery are predicated on the importance of reading connected text, the interconnectedness of reading and writing, and reading as a process. The program resonates with American educators who have recognized a need to stem declining literacy levels (Hiebert, 1995).

Pressley (1995) noted that Reading Recovery stresses five exceedingly explicit concepts fundamental to the reading process: the direction of print moves from left to right, a quick return to the left from the right side of a page ensures continuity of thought,

and cues to meaning appear in illustrations and sentence structure, the reader must self-check for what is sensible, and the reader must reread to achieve clear understanding. Pressley recognized that intensive, repeated practice of such basic strategies might actually be the key to fostering growth among those first graders who were making little progress in a traditional classroom.

Reading Recovery's lesson framework typically includes six to seven instructional activities within a 30-minute session. Each of the six activities commands the student's interaction with text, whether reading, rereading, writing, or cutting sentences apart. The seventh activity involves very basic identification, an interaction with words and letters, but only if deemed necessary (Clay, 1996a). The instructional format earned praise from Pikulski (1995) for providing daily opportunity for rereading recently mastered material.

Hobsbaum (1996) welcomed the concept of scaffold learning, the deliberate teaching of reading skills to empower the student to solve the challenges posed by new literacy tasks embedded in Reading Recovery. Hobsbaum also praised the use of writing as a means of connecting children's prior knowledge to new situations.

Similarly, Barnes (1997) noted that Reading Recovery provided valuable, accumulated knowledge regarding the ways students read and write. Browne (1997) lauded the strong emphasis on the connections between reading and writing. However, some researchers found discrepancies between Reading Recovery and research findings or contradictions between Reading Recovery and aspects of their own practice. For example, Santa and Høien (1999) realized that Reading Recovery assumes that students acquired word recognition skills through reading connected text. Teachers trained in Reading Recovery encourage children to check sentence context or confirm the

prediction of word meaning, rather than look at the interior details of words. Hiebert (1995) noted an emphasis on word recognition and spelling but questioned the omission of emphasis on understanding the tasks and composition writing in what has been described as an integrated program. Barnes observed that Reading Recovery seems to invalidate invented spelling as spelling lessons moved swiftly from word meaning to mechanics in the quest for accuracy.

In an investigation into the progress of three matched groups of first graders, Iversen and Tunmer (1996) found that students in the Reading Recovery group were particularly lacking in phonological-processing skills. The systematic introduction of grapheme-phoneme activities plus word analysis into Reading Recovery lessons accelerated the progress of those students under study.

Pikulski (1995) noticed a similar lack of phonemic awareness, whereas Santa and Høien (1999) suggested that the phonemic component is perhaps emphasized in strategies for writing but, apparently, not for reading. These inconsistencies seemed to reveal a strong programmatic dependence on whole-language tenets (i.e., word-level information comes from context and only incidentally through sounds of letters or lists of words).

Substantial research support has been forwarded for a more than incidental contextual reference to letter-sound correspondence in an alphabetic language such as English during the developmental phases of literacy learning (Adams, 1995; Beech & Harding, 1995; Ehri, 1996; Felton, 1995; Foorman & Liberman, 1996; Forman, Francis, Novy, & Liberman, 1996; Goyen, 2000; Hurford, Schauf, Bunce, Blaich, & Moore, 1995; McGuinness, McGuinness, & Donohue, 1995; Simmons & Kame'enui, 1998; Stanovich & Siegel, 1996; Vandervelden & Siegel, 1995). Gredler (1997) questioned the effective component of the instructional environment itself. With teacher-direct talk serving only

adjunctly to instruction, expression of possible fear of failure, fear of the adult in charge, or fear of the adult's reaction to student error could negatively impact learning. The possibility that lack of verbal interaction could hinder academic progress was raised. These are possible questions that concerned Gredler when he researched the ways in which students would adjust to Reading Recovery instruction and the classroom environment.

### *Staff Development*

Commitment to Reading Recovery includes participation by teachers in a summer orientation just prior to the start of the school year, weekly class sessions, peer observations, and copious record keeping and national conference attendance in addition to the continuous contact required of trained practitioners. Pikulski (1995) identified Reading Recovery as the “most defined and intense consultation” (p. 37) of any intervention program.

Meeting weekly for 2 1/2 hours, Reading Recovery teachers-in-training hone their observational skills by watching and, then, critiquing an actual lesson taught behind a one-way mirror by a peer attending to one of their own students. Prodded by the teacher-leader, 1st-year teachers analyze and interpret the reading and writing behaviors to design and implement an individual program that would merge the idiosyncratic into a theoretical base to support a new teaching mode.

Browne (1997) cited the training course as an opportunity for teachers to increase their observational powers. Spiegel (1998) recommended Reading Recovery for consistently presenting teachers with opportunities to analyze their own instruction critically and to become more involved with their students' progress. Teachers were evaluated for how well they recognized their students' reading problems.

### *Time Allocation*

In a subsequent study, Allington et al. (2000) found that a third of designated remedial time was spent in nonacademic activities. By 2000, Allington and McGill-Franzen (2005) teamed to compare the quantity of reading instruction in Chapter I and special education remediation programs. They concluded that increasing the quantity and quality of homeroom and supplemental reading instruction would serve the eligible population more adequately.

Barr (2000) discovered that students' time, instructional materials, and teacher experience all affected the at-risk students' reading levels and concluded that time allocation for low-achieving students might be the easiest to adjust. Meanwhile, Clay (1996b) introduced Reading Recovery as a means to accelerate the progress of children who had made little progress after a year of formal schooling. Clay (1996b) believed that, because of Reading Recovery's 30-minute lesson format, a "child must never engage in unnecessary activities because that wastes learning time" (p. 9). Prime among other distinctive components, the format featured active learner engagement with text in carefully structured time segments.

The rigorous adherence to time on task caused Barnes (1997) to question students' lack of time to reflect on their learning during lessons. Browne (1997) suggested metacognition was not a Reading Recovery goal for its individualized lessons but, rather "reading for meaning . . . to access information" (p. 295). Rasinski (1995) questioned not the concept of managed learning time but the practice of using it in research studies purporting to contrast the characteristics of comparable reading interventions. All but Reading Recovery offered less focused instructional time, and this asserted that a student staying on task was very important for the success of the Reading

Recovery program.

*Student Selection*

Reading Recovery was started as a safety net for New Zealand children who, after a year of schooling, had made negligible progress by 6 years of age. The selection process in the United States has two additional complexities. First, American 6-year-olds have not commonly experienced a year of schooling because kindergarten is not mandatory nationwide. Not all American children will have experienced a year of reading instruction either because all kindergartens are not academic. Second, Reading Recovery teachers administer the Reading Recovery Observational Record's battery of six measures individually, not just to those who make negligible progress but to any portion of up to and including an entire first-grade class in accordance with school district philosophy and resources. Combining the judgment of classroom teachers with results from the Reading Recovery Observational Record, Reading Recovery teachers selected the lowest achieving children as program participants. Shanahan and Barr (1995) reported that "this is a relative notion of at risk rather than an absolute (p. 962) because selection occurs not necessarily from the lowest quartile or quintile of national achievement levels on a text but from individual performance levels relative to peers within each first-grade classroom as assessed by teacher observations.

Mindful of Reading Recovery design, Ohio State University (1992) stated, "To help the lowest-achieving first-grade children" (p. 1). Hiebert (1995) studied Ohio students' entry levels from 1986 to 1991 and found that the mean score for all below average participants during that period placed them in the bottom 35%, rather than in the lowest 0% to 20% achievement band.

Hiebert (1995) also noticed that student names repeatedly disappeared from class

rolls during the introductory 10 sessions as well as between Lessons 10 and 59. Although some participants may have reached the average band early, moved away from the district, or experienced insufficient time for program completion, others were dismissed due to poor attendance or were referred to special education. Yet, Shanahan and Barr (1995) reported that Reading Recovery participation was inclusive without regard to ability or disability because the diagnostic emphases were on only reading and writing competencies. Hiebert and Rasinski (1995) questioned the ethical appropriateness of a program targeting the subset at the bottom of a distribution.

#### *Program Effectiveness and Student Performance*

To evaluate the effectiveness of any program, the purpose, sampling, instruments, methods, analysis, and conclusions must be detailed. To discern a program's effect on student performance, a number of other interactive factors, including demographics, teacher effects (experience, instructional philosophy, etc.), text provided (design selection and administration), and data collection and interpretation, must be considered over the short- and long-term goals.

Since the American inception of Reading Recovery, many published studies reported impressive, short-term outcomes. Available analyses of long-term studies are few, but they are increasing in number as cohorts of Reading Recovery participants advance through elementary schools and provide test data.

Initially, researchers were predominantly those such as Pinnell, Lyons, Deford, Bryk, and Seltzer (1995) who were directly involved with Reading Recovery at its primary training sites in Ohio, Illinois, and Texas because they had liberal access to data compiled and stored at Ohio State University. Results for all studies from 1988 to 1994 were positive and supportive of Reading Recovery.

In the past 5 to 6 years, Iversen and Tunmer (1996), Hiebert (1995), Shanahan and Barr (1995), and Center (1995) published reviews. Some of them relied upon data from the in-house Reading Recovery collection system, whereas others generated precise data from independent measures of ability of children to read. Overall, results were inconclusive, restricted by incomplete data, flaws in sampling procedures, and failure to control for regression to the mean that limited the scope of support for one another's data.

Reutzel (1999) found that Reading Recovery was not effective with all children, but was "effective with some, even many" (p. 97). Ross (1995) suggested that Reading Recovery might, perhaps, best suit schools with relatively few students at risk of failure because of its focus on so few children. It is important for teachers to have access to the most effective practices that target children who are at risk of failure because of the influence of student variables. Reading Recovery serves the lowest achieving students. The lowest achieving students in seventh grade, without exception, are selected to receive the program.

### *Summary*

The literature analysis presented details addressing why many children come to school not quite ready to participate in reading instruction. The physical welfare of the child, socioeconomic status, and environmental factors all contribute to a child's ability to learn how to read. School districts nationwide employ an assortment of programs in efforts to arrange intervention methods for students performing below their classmates in reading. Higher quality interventions are more triumphant when used at the beginning of a child's education.

Reading Recovery is a remediation technique created by Clay (1996b), and it initially began in her country. This intervention strategy involves first-grade students and

is based upon Clay's (1995a, 1995b, 1996a, 1996b, 1998) research in which she detailed how children learn to read. Although the program has many who oppose it, the Reading Recovery program has grown in usage over the last 2 decades in the United States.

Pinnell (2005) contended, vehemently, that Reading Recovery yields discernible changes in the growth of the lowest performing first-grade students who, once finished from the program, read at an average level when compared to their peers in the same class.

Research points to Reading Recovery as an intervention program with promise.

Children discontinued with the program are better readers than those who were eligible but who did not receive the services. Research also indicates that the gains made by first graders of the program gradually lose the gains that were made. According to current literature, to help children sustain the reading gains that they make, it may be wise to incorporate Chapter I resources as well. Torgensen (1998) argued that the effectiveness of Reading Recovery could be improved substantially if it provided more explicit and systematic instruction in the alphabetic principle (phonics). There are also other independent analyses showing that Reading Recovery does not accomplish the goal of preventing reading difficulties in young children as effectively as its in house research tends to show nor has it responded to research identifying its weaknesses (Elbaum, Vaughn, Hughes, & Moody, 2000).

### Chapter 3: Methodology

This chapter describes the populations examined and the measurement instruments, data-collection procedures, and data analyses used in this study. The Reading Recovery program includes not only an instructional program but also its own evaluation system that aligns with programs. Most data cited regarding the effectiveness of Reading Recovery program are gathered through the Reading Recovery program evaluation system. This system uses a unique pre- and posttest research design. The measures and the program, along with data-collection procedures that are controlled within the Reading Recovery program implementation system, create an increased potential for bias in the result of an evaluation. Because most of the data available regarding the success of Reading Recovery program come from its own evaluation system, the research design and the measures used in this system are discussed first.

#### *Research Design*

The participants of this research study included those students recommended for Reading Recovery intervention during the seventh grade of their schooling. This study was designed to determine whether Reading Recovery helped to sustain reading achievement 3 years after the experiment. Two types of data were used to compare the performance of the Reading Recovery children with that of the comparison group. First, the achievement of the comparison group is used to establish a band of achievement. The band is a half standard deviation above and below the mean in each of the areas taught to the Reading Recovery students and measured by the Clay diagnostic measures. If a Reading Recovery student's scores end up within this band, then, the child is considered successful and is discontinued (Pikulski, 1995). Secondly, the data are analyzed and compared to the pre- and post test gains made by the Reading Recovery children

with the comparison group to see if the children in Reading Recovery gained at a faster than normal rate while in Reading Recovery.

### *Populations*

The study considered students who had below FCAT scores were identified as eligible for Reading Recovery in a year and eligible for Chapter I in 2 years. Those participants attended one of eight alternative schools. No school had more than 230 students or fewer than 175. Four schools had an enrollment of 200 to 220 students, two schools enrolled 180 to 195 students, and two schools enrolled 150 to 165 students.

In the 1st year, each school qualified for the federally funded, supplemental language arts program known as Chapter I. Based on the number of students receiving free and reduced-price lunch, Chapter I supported at least one Reading Recovery teacher and one reading specialist in every elementary school. Thus, county schools provided resources that included one reading specialist (.5 Reading Recovery, .5 Chapter I) at a minimum; larger schools had as many as three reading personnel.

### *Designation of Groups*

The school district had 140 students who qualified for supplemental instruction during the fall term. Of this group, 73 were male (55%), and 67 were female (45%). One hundred and sixty students (97 males and 63 females) entered the Reading Recovery program; 20 males and 20 females remained on a waiting list.

In 2006, all seventh graders in the school district took the Stanford 9 Achievement Test. Thus, 105 seventh graders (66 males and 39 females) who had participated in Reading Recovery, 106 seventh graders (53 males and 53 females) who had Chapter I instruction for 3 years, and 25 eligible seventh graders (15 males and 10 females) who had received no supplemental reading instruction had scores from the vocabulary and

reading comprehension subtests of the Total Reading section of the Total and Complete Battery.

During the 2nd year of this study, 27 documented subjects moved out of the county, 19 were retained one grade level, and 15 could not be located because they may have acquired a special education label or moved away without requesting official records. No Stanford 9 information was available for these 77 students missing from the ninth-grade roster.

### *Reading Instruction*

All seventh graders received regular, daily classroom instruction or two classroom sessions per week to the whole class as a group, in small literature groups, or in some combination of these. Additionally, Reading Recovery and Chapter I students received 30 minutes extra instructional time per day for reading-related activities with a literacy specialist. Classroom teachers held bachelor's or master's degrees, whereas all 37 reading teachers in the school district held master's degrees and the reading specialist endorsement on their postgraduate professional license. Reading specialists had classroom experience ranging from 4 to 24 years.

### *Selection Process*

Student eligibility for supplemental remedial reading services in the county was dependent on input from classroom teachers and reading specialists working in tandem to identify children most in need of increased reading support. With assistance from Reading Recovery teachers, seventh-grade teachers used screening devices as diverse as identification of alphabet letters and the Metropolitan Reading Test to assess all seventh graders. From these collective, informal assessments of reading achievement, the teachers ranked all students in their classrooms.

Reading Recovery teachers, using Clay's (1996a) diagnostic survey, then, retested the bottom half of each ninth-grade class and ranked the students according to their scores on the survey's sentence dictation test. The four lowest scoring students in the seventh grade became Reading Recovery entrants.

### *Program Description*

Supplemental reading programs operational in the school district featured decoding, encoding, and comprehension components, but they occurred under differing conditions in the classroom.

### *Reading Recovery*

Reading Recovery lessons emphasized reading and writing in daily, one-on-one tutoring sessions that followed a prescribed format (Center, 1995). Reading Recovery is different from other reading programs because an at-risk child in Reading Recovery receives individualized help based on his or her strengths identified through careful observations of the child's reading and writing behaviors. A specifically trained teacher works with only one child at a time for 30 minutes daily. A typical reading program uses a group setting with one teacher working with several children at once.

Additional unique features appeals to visual, aural, and kinesthetic senses and included a collection of short trade books (graded according to difficulty) to be read and reread by the student, a large pad of blank paper, pencils, and a tray with magnetic letters.

Instructional program procedures for Reading Recovery were carefully scripted to promote consistency and ensure quality (e.g., teachers would engage a student in sentence composition with close monitoring for accuracy and would record a daily analysis of student progress). The buzzing of a kitchen timer signaled the end of one activity and the beginning of another.

Students at four instructional reading levels (readiness, primer, Grade 7, and Grade 8) in Reading Recovery's program had several assorted age-appropriate books available for daily reading. The actual amount of text read daily by each student varied by book level, but, due to close program monitoring, it is safe to assume that students read a minimum of the three books prescribed per session.

Progress through book levels from 1 to 20 became a tangible indicator of student success. The short trade series took the students from decoding 19 words grouped in phrases (e.g., "great pie," "good looking," and "How are you doing?") at Level I to as many as 525 words grouped into compound and complex sentences (e.g., "When Joe ran into his room, some children began to laugh") at Level 20.

Instruction continued until students consistently demonstrated competence in actively creating meaning by reading text at approximately 90% accuracy or above and by demonstrating strategies providing a measure of independence in reading. At this juncture, the Reading Recovery teacher discontinued the Reading Recovery sessions, and the students joined the average band of readers in class and receive in-class, group instruction with them.

Although the approximate time for recovery to median ninth-grade level and a measure of independence in reading is 12 to 14 weeks or 60 lessons (Clay, 1998), some students exited the program more quickly, and some required more than the minimum number of lessons.

### *Chapter I*

A highly individualized, eclectic combination of whole language, phonics, and guided reading philosophies determined content in each Chapter I class. Lessons provided reading and writing opportunities, but the specific design of the instructional

format within the 30-minute block of time relied upon the creativity of the reading specialist.

Instruction started with word analysis, proceeded to predicting content, moved to using strategies in oral or silent reading of text, and concluded with revision of predictions based upon an individual's comprehension and whole-group discussion. Typical activities included shared, silent, and choral reading; readers' theater; skill lessons in workbook or on computers; response journals; and creative writing. Materials included basals, trade books, and poetry, often enriched with art, music, and cooking extensions.

Chapter I teachers conducted daily small-group instruction. The format followed site-based decisions regarding program models. Some schools preferred an in-class model; others supported pullout programs. Small groups averaged six students per class, but the number of students' inclusion models varied depending upon the planned activity.

The actual amount of time Chapter I students were involved with text varied widely by day, by instructional model, and by teacher. From as much as optional 25 minutes on a traditional day in a pullout program or 45 minutes to 1 hour in an inclusive setting, students could read as little or not at all, depending on the day's planned activities. It is safe to assume that Chapter I students overall read more text at their instructional levels than they would have read in their homeroom class, but the approximate amount of text read cannot be ascertained due to an absence of standardized resources, curriculum, and supervision. Students continued in Chapter I classes for the entire school year unless they demonstrated accelerated progress in decoding or comprehension consistently throughout a 6-week marking period. At that point, the Chapter I teacher could, in consultation with the classroom teacher and parents, exit any

student upon successful completion of an informal reading inventory posttest administered one-on-one by the Chapter I reading teacher.

In the absence of evidence of sudden, major improvements, reading specialists relied upon a combination of simple gains in posttest informal reading inventory scores at year end, classroom teacher and specialist assessment of decoding and comprehension skills, and standardized test scores when applicable (Grade 9) to determine whether a student's independent reading level was commensurate with the class average to ensure continued progress without specialized support.

Because the composition of homeroom classes changed annually, the average independent reading level in homerooms also changed as teachers tested students and ranked their reading achievement each year. Thus, it is possible and probable that some Chapter I students had already spent years in the program when this study was carried out.

#### *Measurement Instruments*

The Clay (1996a) diagnostic survey's six proficiency indicators were grouped together to profile of the language competencies of very young school children as a pretest. The first 4 tasks assessed learning in areas with finite sets of information such as (a) concepts about print, (b) letter identification, (c) sentence dictation, and (d) word test. The remaining two subtests, text reading and writing vocabulary, were scored differently.

Raw scores identified language strengths and indicated initial text level for instructional purpose or vertical advanced text level for instructional purpose or verified advanced text level for discontinuation. Standard directions for introducing and coaching each test accompanied the survey.

Three of the six survey subtests (sentence dictation, word test, and text reading)

accurately revealed the acquisition of reading behaviors, rather than those that evaluated familiarity with several conventions of print (concepts about print, letter identification, and writing vocabulary). Reading Recovery teachers used those first 3 subtests for posttest analysis at the discontinuation of Reading Recovery sessions as well as at the end of the school year. Alternative word lists and dictation sentences plus a variety of trade books provided fresh, accurate testing materials. Reading Recovery teachers scored their own diagnostic surveys by hand from the results of a pretest. In order to eliminate any possible scoring bias, Reading Recovery colleagues from neighboring schools in the school district administered and hand scored each other's posttest. The Reading Recovery teacher leader reviewed the scores compiled by Reading Recovery teachers for each subtest of every student's diagnostic survey.

All students included in the applied dissertation took the Reading Recovery diagnostic survey pretest as ninth graders. Students who had experienced Reading Recovery instruction during their ninth-grade year participated in posttesting.

#### *Norms*

Research norms for Clay's (1996a) diagnostic survey were developed using two samples of New Zealand students from five different urban schools. The 1968 group included 320 children aged 5-7 years; the 1978 group included 282 children aged 6-7 years.

#### *Reliability*

Clay (1995b) reported that internal consistency analysis was based on groups of children ranging in size from 34 (writing vocabulary) to 100 (concept of print, letter identification, and word test). Test-retest reliability ranged from .73-.97, split-half ranged from .84-.97, and Cronbach alpha ranged from .83-.97. These ranges of coefficients

indicated that the survey's subtest measured literacy competencies consistently.

### *Validity*

To illustrate how to increase test scores on three survey subtest, Clay (1995b) reported that criterion-based validity data correlated studies with the Reading Recovery program. Concepts of print, letter identification, and word test relate to increase in children's reading of frequently occurring words. The validity coefficients, ranging from .79-.85, were high. That suggested that the students in the sample produced a broad range of scores on the skills being measured.

### *Classroom Reading Inventory*

Silvaroli's (1997) classroom reading inventory is an informal reading inventory that enables an instructor to identify specific word-recognition errors and approximate a program entry-level through oral reading of graded word lists and graded paragraphs. Approximate independent, instructional, and frustration reading levels can be determined through the use of the answer key provided. Examples of five common student errors serve as examiner guidelines in the manual, but reading specialists are encouraged to evaluate and judge whether errors significantly hinder students' understanding or fluency. Thus, scoring by Chapter I teachers at the building level retained a subjective element because students frequently would offer synonyms, examples, or anecdotes.

Neither reliability nor validity information were available for the classroom reading inventory (eighth edition), but there was a review of the fourth edition by Conoley and Kramer (1998). Two independent reviewers reported that the graded word lists and oral paragraphs were appropriate in differences from level, the manual had clear directions for administration in differences from level, the manual had clear directions for independent reviewers reported that the graded word lists and oral paragraphs were

appropriate in differences from level, and the manual had clear directions for scoring.

*The Stanford 9 Achievement Test*

According to Worthan (1995), standardized tests “provided samples of behavior that can be served and evaluated according to an established standard” (p. 50). The Stanford 9 can provide informal evaluation measures and teacher observations with, as stated by Worthan, “uniform procedures for administration, quantifiable scores, norm referencing validity, and reliability” (p. 50).

Although standardized test can provide a school district with periodic reviews of general educational progress, they also can certify students’ competence. Test scores frequently help school personnel to make student placement decisions by determining eligibility for specialized programs and monitor and evaluate the progress of specific groups by comparing scores across grade levels and across years.

The Stanford 9 included two reading subtests, vocabulary and reading comprehension, in its total reading composite. The vocabulary subtest consisted of 20 multiple-choice items focusing on word knowledge where students would identify synonyms, use context clues to determine word meaning, and explore multiple meanings of a word within different contexts. The reading comprehension subtest consisted of 30 multiple-choice items focusing on student understanding of the three types of reading selections: recreational, textual, and functional. Raw scores of each subtest were summed to create the total reading scores for analysis in this applied dissertation. Although a discrepancy may exist between what tests measure and current instructional trends in reading, reading scores can be used to chart development progress. Word study cannot describe a student’s integrated system of reading behaviors (Clay, 1996b), but it can reveal the extent to which a student understands components of language. Similarly,

reading comprehension cannot identify how effectively a student constructs meaning, but it can provide multiple opportunities for a student to read continuous text, recall information, and demonstrate understanding.

The Stanford 9 total reading scores are a useful measure in this study because test data, however imperfect, drive curricular decision making and policy implementation. Reviewers of the Stanford 9 concurred that the ninth edition was truly different from previous editions with attention paid during test creation to content of leading textbook series, educational trends, and standards of national professional organizations in content areas.

Although new items and performance standards were included to reflect changing school curricula, neither development was relevant to the total reading component. The Stanford 9 noted that the Kuder-Richard Formula 20 reliability coefficient was used to estimate the internal consistency of the multiple-abbreviated batteries for the standardization samples norm in Spring 2002 and used in this applied dissertation. Pertinent data for Grade 9 includes total reading ( $r = .92$ ,  $SEM = 2.89$ ), reading vocabulary ( $r = .81$ ,  $SEM = 1.77$ ), and reading comprehension ( $r = .88$ ,  $SEM = 2.26$ ). An alternative forms coefficient reported for equating forms SA and TA for the multiple-choice assessment was  $r = .89$  and  $SEM = 4.90$ .

The technical data report authors reported validity data from item difficulty values, scaled score demonstrating annual growth, correlation studies between assessment forms, adjacent grade-level subtest, and test editions. The authors provided criterion and construct validity but suggested that detailed comparison studies of test content and curriculum objectives of specific school systems contemplating use of Stanford 9 should provide the content validity. Stanford 9 was machine scored outside of

Palm Beach County, and scores were reported on printouts for the school system.

### *Data Collection*

The researcher collected data for the applied dissertation in two stages. Prior to data collection in Fall 2004 and Winter 2005, two letters detailing and refining the purpose and design of this study were sent for review to the school district administrator and the committee in charge of research. The first letter attained permission to visit schools to review class rolls as well as produce a master list of successfully discontinued Reading Recovery students from the 2005-2006 school year who were identified by gender and school. Initially, the researcher gained access to each of the four alternative schools by phoning the principal and briefly explaining the nature of the study. The researcher drove to each school, proffered the official letter of support and permission from the reading supervisor, and collected ninth-grade Stanford 9 scores from 2005 for the population cited above.

A number of successful Reading Recovery participants had moved within and outside the county between Grades 9 and 10. Tracking the in-county students was time-consuming, and it frequently produced inconclusive results. Some students had transferred to private schools; some had chosen home schooling. Sixteen students had been retained and had not taken any of the Stanford 9 tests.

To create a larger, more inclusive study, the researcher expanded the Reading Recovery group to include all Reading Recovery students, rather than merely successful students who had participated in any Reading Recovery lessons. The study was further enlarged to include all the students served by Chapter I or placed on its waiting list to serve as additional comparison groups.

Following receipt of the second letter, the reading supervisor and Reading

Recovery teacher-leader met with the researcher to discuss their particular concerns about objectivity in the study. The researcher reviewed all the files for the seventh-grade Stanford 9 scores for students eligible as seventh graders for Reading Recovery, Chapter I, or placed on a waiting list for remediation. All data used in the applied dissertation were systematically collected from preexisting school or Reading Recovery files over a 10-month period.

#### *Data Analysis and Procedures*

In this applied dissertation, reading development, academic achievement, and self-esteem were evaluated using Stanford 9 test scores to describe the long-term reading achievement of students who received Reading Recovery instruction in the spring. The following research question was investigated: What will be the effectiveness of a Reading Recovery program on improving the FCAT scores of at-risk students?

Because random assignment to Reading Recovery groups was not possible, this study analyzed means and standard deviations to compare the relative performance of three major populations: Reading Recovery, Chapter I, and wait listed. Descriptive statistics for the three populations were based on a total of three quantitative variables obtained for each Reading Recovery population in the study, on two quantitative variables for the Chapter I population, and on one for the wait-listed group. Raw scores used for all statistical analyses of Reading Recovery populations included raw scores used for the statistical analyses of the FCAT.

## Chapter 4: Data Analysis

This chapter reports the results of the analysis of the data. This section includes tables that describe characteristics of the populations and presents the results of the analytical procedures.

### *Profile of the Populations*

Table 1 presents the number and percentage of students eligible for reading support during the fall of 2006. The 230 children represented the lowest achieving of all seventh-grade students screened in the county. The 160 lowest scoring children entered the Reading Recovery program, the next higher scoring group of 113 participated in Chapter I, and the 40 highest scoring students in the group of 313 became the waiting list. Thus, half of all eligible students participated in Reading Recovery program, and a little more than a third of all eligible students had Chapter I instruction.

Table 1

### *Student Participation in Supplemental Reading Programs in the 2006-2007 School Year*

Group	<i>N</i>	% of total
Reading Recovery	160	51
Chapter I	113	36
Wait listed	40	13

### *Attrition*

Table 2 profiles the attrition during the 2-year period subsequent to Grade 8 of subjects in the applied dissertation. Fifty-six subjects (18%) from the original population

who had participated in the supplemental reading instructional programs could not be included in the final analysis of scores. They had been retained a grade or had moved away from the county. No record existed for the remaining 21 students (7%) who had also received instruction in eighth grade. Thus, one quarter of the original group of eligible students was unable to participate in the follow-up study.

Table 2

*Attrition of Program Participants  
in the 2006-2007 School Year*

Condition	<i>N</i>	% of total
Moved	37	12
Retained	19	6
Unknown	21	7

*Note.* There were 33 possible participants.

*Total Number of Student Participants*

The numbers of participants eligible for inclusion in the 2.5-year study and the corresponding percentage of the total for each instructional group appear in Table 3. The 39 students in the Reading Recovery group (16%) included those students who only had Reading Recovery in Grade 8.

The 105 students in the Chapter I group (45%) included all eighth-grade students in the lowest 25% of their homerooms who were ranked just above Reading Recovery students and participants in Chapter I instruction. The 25 students in the waiting list (11%) identified those students eligible but did not need instruction because of relatively

high scores in the bottom half of the class.

Table 3

*Student Eligibility for Inclusion in Study in 2006*

Group	<i>N</i>	% of total
Reading Recovery	39	16
Chapter I	105	45
Wait listed	25	39

*Study Participants*

The 165 participants in the 2006-2007 school year follow-up study is delineated in Table 4. The number of students in Reading Recovery programs remained constant at 105 students (63%).

However, from the 131 students receiving Chapter I (106 students) and those wait listed (25 students), the researcher selected 60 students (30 males and 30 females) using a table of random numbers to form a comparison group. Fifty-three of the 60 students (88%) in this group had experienced Chapter I instruction, whereas 7 students (12%) had not participated in supplemental reading programs of any kind. Although it can be assumed that test scores for the large populations were representative, scores for the smallest populations in the study might be somewhat extreme or less typical of the total wait-listed population.

*Results*

Descriptive statistics for the third populations were computed based on a total of

three quantitative variables obtained for each Reading Recovery population in the study and on two quantitative variables for the comparison group. One hundred and five students (64%) in two of the populations under study had Reading Recovery instruction. The average raw score for the group of 39 students who had only Reading Recovery was 27.4, which was 2.2 points higher than the average raw score.

Table 4

*Student Participation in Study in 2007*

Group	<i>N</i>	% of total
Reading Recovery	39	23
Chapter I	53	32
Wait listed	7	4

Reading specialists then administered the Reading Recovery program to all remaining ninth graders positioned in the bottom half of the group produced by classroom teachers for the Reading Recovery program. The specialist collated results to identify the students in the lowest 25% and placed them on an eligibility list to become the Chapter I roster. Once the reading specialist had scheduled the six neediest students per half hour Chapter I class (whether a pullout or in-class model), any students remaining who were not served by reading specialists composed a waiting list. Students on the waiting list received classroom reading instruction only.

Raw scores used for the statistical analyses of the Chapter I population included the number of years spent in the program and the Stanford 9 test scores. The wait-listed population had only Stanford 9 test scores to consider. Research showed that a large

percentage of children who were the poorest readers in their schools after 1 year responded quickly to the Reading Recovery program.

*Summary of the Test Populations*

Means and standard deviations for each population are presented in Table 5. The average scores for the three populations in the study are within 1 to 6 points of each other. The performances of each of the three populations on the Stanford 9 were very similar.

Table 5

*Summary of Florida Comprehensive Assessment Test Scores: Three Populations*

Group	<i>N</i>	<i>M</i>	Range	<i>SD</i>
Reading Recovery	39	27.4	7-44	9.5
Chapter I	53	28.3	11-45	8.6
Wait listed	7	31.1	14-46	10.9

The two populations with no Reading Recovery instruction had the highest mean scores. The average raw score for Chapter I was 28.3, whereas the wait-listed group had a raw score of 31.1.

Missing from the three sample population totals are statistics of some interesting subgroups formed by considering two additional independent variables. These variables displayed the performance of three large test populations and revealed observed differences.

The first variable was whether or not a Reading Recovery student reached one of

Clay's (1998) criteria called discontinuation. At the point when students would attain the Reading Recovery program goal of successfully reading and writing within the average band or slightly above the average of their homeroom class (trade book Level 16 or 18), Reading Recovery teachers would discontinue Reading Recovery instruction. Because discontinuation depended to some extent upon the number of lessons provided, the second variable was whether a student met his or her reading goal in the optimal intervention period of 12 weeks or 60 lessons.

Table 6 depicts discontinuation data for the two Reading Recovery populations under study. A clear demarcation existed between those who achieved discontinuation and those who did not. For those who attained this benchmark, the conditions under which success occurred are noted. Table 7 adds Chapter I details to the summary of discontinuation and lesson data for Reading Recovery populations.

#### *Characteristics of Reading Recovery Students*

Sixteen of 27 students completed the program in fewer than 61 lessons. On average, they received a raw score of 32.3 of 50 items on the total reading test of the Stanford 9. This was the second highest scoring group in the study population. It was slightly higher than the means of three other groups: (a) all Reading Recovery Chapter I students (mean = 31.6) who had fewer than 61 lessons with 1 year of Chapter I (see Table 8), (b) the Chapter I only group (mean = 31.3) with 1 year of instruction (see Table 8), and (c) the number of years in the Reading Recovery program.

The remaining 11 Reading Recovery only students (mean = 24.1) achieved discontinuation with more than 60 lessons. Their mean was slightly less than the nondiscontinued group (mean = 24.9) completing the program in more than 60 lessons in the Reading Recovery program. The nondiscontinued group correlated with the mean of

all Reading Recovery and Chapter I students who also had more than 60 lessons plus 1 year of Chapter I (mean = 24.9; see Table 7).

Table 6

*Summary of Florida Comprehensive Assessment Test Scores:  
Reading Recovery Populations*

Group	<i>N</i>	<i>M</i>	Range	<i>SD</i>	%
Reading Recovery					
Discontinued					
61 lessons	16	32.3	16-44	7.7	41
60 lessons	11	24.1	15-44	9.1	28
Not discontinued					
61 lessons	3	21.3	7-37	15.0	8
60 lessons	9	24.9	12-40	8.8	23
Reading Recovery and Chapter I					
Discontinued					
61 lessons	25	30.5	14-45	8.5	38
60 lessons	13	26.2	16-35	5.9	20
Not discontinued					
61 lessons	20	20.5	5-36	8.7	30
60 lessons	8	19.1	7-31	7.3	12

Twelve students (31%) did not attain discontinuation in any number of lessons.

Nine of these students had a mean score of 24.9 with more than 60 lessons.

The remaining three nondiscontinued students had, with fewer than 61 lessons, a mean score of 21.3. The nondiscontinued students matched the average score (mean = 21.3) of all Reading Recovery and Chapter I students who had more than 60 lessons and 2 years of Reading Recovery (see Table 7). This was only slightly lower than the score

(mean = 21.8) of the Chapter I only group with 2 years of Reading Recovery instruction.

Table 7

*Details of Florida Comprehensive Assessment Test Scores:  
Number of Years in Reading Recovery*

Years	Fewer than 61 lessons			More than 60 lessons		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
0	19	30.6	9.5	20	24.5	8.7
1	19	31.6	8.8	13	24.9	6.9
2	20	23.0	8.0	8	21.3	8.0
3	6	18.3	9.4	--	--	--

*Note.* Zero years in Chapter I identify the Reading Recovery only population.

The data indicated that Reading Recovery students who were successfully discontinued from the Reading Recovery program in the optimal time frame of fewer than 61 lessons scored higher than similar groups in this study. It also appeared that Reading Recovery students who received more than 60 lessons had means very similar to each other, whether discontinued or not.

#### *Characteristic of Reading Recovery and Chapter I Students*

Sixty-six students received Reading Recovery instruction in combination with Chapter I. Of these, 38 students (58%) from this population were successfully discontinued from the program as delineated in Table 7.

Twenty-five of 38 students completed the program in 61 lessons or less. On average, this group received a raw score of 30.5 of 50 items on the FCAT. This raw score was virtually identical to the average score (mean = 30.6) of all Reading Recovery only students who had fewer than 61 lessons (see Table 7).

Table 8

*Summary of Florida Comprehensive Assessment  
Test Scores: Comparison Populations*

Group	<i>N</i>	<i>M</i>	Range	<i>SD</i>	%
Chapter I					
1 year	38	31.3	14-45	6.0	63
2 years	12	21.8	12-42	10.7	20
3 years	3	19.7	11-26	6.3	5
Wait listed	7	31.1	14-46	10.9	17

Within this same group of 25 students, a subgroup of 14 also had received 1 year of Chapter I instruction and averaged 34.3 raw score points (a) the Chapter I only group (mean = 31.6) with 1 year of instruction (see Table 7), (b) the Reading Recovery population (mean = 27.4), and (c) the number of years in the Reading Recovery group (mean = 31.6; see Table 7).

A second subgroup of eight discontinued students with fewer than 61 lessons had 2 years of Chapter I instruction (mean = 26.5). Their score exceeded the average scores (mean = 23.0) of the total Reading Recovery population with 2 years of Chapter I (see Table 7).

Three students in the third subgroup of discontinued students with fewer than 61 lessons had 3 years of Chapter I (mean = 23.3). This score exceeded the average score

(mean = 19.7) of Chapter I students with 3 years of Reading Recovery instruction. The average score of the total Reading Recovery and Chapter I population (mean = 18.3) with fewer than 61 lessons (see Table 7).

Reading Recovery was the only intervention that showed long-term improvements in reading. At the end of the 70 days of instruction, Reading Recovery children were reading five levels ahead of children who received regular remedial reading lessons. Although (in contrast to Reading Recovery children) the control group continued to receive lessons for the rest of the year, Reading Recovery children were still three reading levels above the remedial group average when all children were tested the following year.

The 13 remaining students (mean = 26.2) achieved discontinuation in more than 60 lessons. This mean was higher than the score (mean = 24.1) for Reading Recovery students only who discontinued with more than 60 lessons and only slightly lower than the average score (mean = 26.5) for Reading and Chapter I students who had fewer than 61 lessons with 2 years of Chapter I (see Table 9). Nine of 13 discontinued students with more than 60 lessons (mean = 25.6) had 1 year of Chapter I, but this score was lower than the average score (mean = 28.9) for all Reading Recovery and Chapter I students with 1 year of supplemental instruction.

A small discontinued subgroup of four students received 2 years of Chapter I and averaged 21.8 raw score points (see Table 9). It was the highest scoring group among students who had 2 years of Chapter I. This score exceeded the average mean for all Reading Recovery and Chapter I students with 2 years of instruction (see Table 8).

Reading Recovery and Chapter I students who were discontinued from the Reading Recovery program after three or more lessons scored consistently higher with 1 to 3 years of Chapter I support when compared to similar groups in this study. Nine of

these had 1 year of Chapter I. Five students with fewer than 61 lessons received a raw score of 24.2 of 50 items on the FCAT. This score was virtually identical to the Reading Recovery only group (mean = 24.1) that discontinued with more than 60 lessons. The remaining four students with more than 60 lessons received a raw score of 23.5. This raw score was slightly higher than the average score (mean = 23.0) for all Reading Recovery and Chapter I students with 2 years of Chapter I and fewer than 61 lessons.

Table 9

*Details of Florida Comprehensive Assessment Test Scores: Number of Years in Chapter I*

Category	Years		
	1	2	3
Discontinued < 61 lessons			
<i>N</i>	14	8	3
<i>M</i>	34.3	26.5	23.3
<i>SD</i>	7.5	7.7	9.2
Discontinued > 60 lessons			
<i>N</i>	9	4	--
<i>M</i>	25.6	21.8	--
<i>SD</i>	7.2	3.0	--
Not discontinued < 61 lessons			
<i>N</i>	5	12	3
<i>M</i>	24.2	20.6	13.3
<i>SD</i>	8.3	7.5	7.6
Not discontinued > 60 lessons			
<i>N</i>	4	4	--
<i>M</i>	23.5	14.8	--
<i>SD</i>	7.0	5.3	--

Twelve of 28 students with a mean of 20.6 and 4 students with a mean of 14.8 had 2 years of Chapter I. Three students with a mean of 13.3 who were not discontinued had 3 years of Chapter I.

The data indicated that Reading Recovery and Chapter I students who were not successfully discontinued from the Reading Recovery program but had 1 year of Chapter I instruction had scores that were similar to other groups in this study. The nondiscontinued Reading Recovery and Chapter I students scored lower than similar groups in this study, despite 2 or 3 years of Chapter I instruction.

#### *Characteristics of Chapter I Students*

One hundred and thirty-one students eligible for supplemental reading services had not qualified for Reading Recovery but had qualified for Chapter I. From this population, the researcher selected 60 students (46%) to serve as a comparison group for this study (see Table 8).

Fifty-three of 60 students (88%) received 1 to 3 years of Chapter I instruction. Thirty-eight of 53 students (72%) in the Chapter I population received 1 year of Chapter I instruction and, on average, earned a raw score of 31.3 of 50 items on the FCAT. This score was the third highest in the study population. The raw score was slightly less than the average score (mean = 31.6) of all Reading Recovery and Chapter I students who discontinued in fewer than 61 lessons with 1 year of Chapter I (see Table 7).

Twelve students (23%) from the Chapter I population with 2 years of instruction earned a mean score of 21.8 (see Table 8). This score nearly matched the average score (mean = 21.3) of Reading Recovery students not discontinued and with fewer than 61 lessons (see Table 6).

The data seemed to suggest that, with a year of instruction, Chapter I students in

this study may score slightly higher than the average for all Reading Recovery students when discontinued with less than 61 lessons and virtually equal to the average for Reading Recovery and Chapter I students with 1 year of Chapter I (see Table 9).

#### *Characteristics of Wait-Listed Students*

The small population of seven students who had neither Reading Recovery nor Chapter I from 2005 to 2006 earned an average raw score of 34.3 of 50 items on the FCAT (see Table 9). This score nearly matched the average mean scores of two groups from two different populations. The Reading Recovery and Chapter I group who discontinued in fewer than 61 lessons had a mean of 24.2 (see Table 9), and the Chapter I only group who had 1 year of instruction had a mean of 31.3.

Although the population was very small (7 students) and some of the scores could be extreme, the data seemed to suggest that students on the high end of the eighth-grade eligibility list from reading support in this study scored equally as well as specific subgroups of students from the Reading Recovery and Chapter I populations. Most of these students made progress in learning to read and write. However, as indicated by the data, these students continue to need long-term support in reading. The data suggested that, although the population was small, students continued to show significant improvement in reading and comprehension skills.

Large-scale and local investigations demonstrated that Reading Recovery is a particularly effective method for correcting the reading difficulties of at-risk children. Clay (1995b) stated, "Rapid progress of these children can be directed for long term evaluation" (p. 22).

## Chapter 5: Discussion, Implications, Recommendations, and Conclusion

This study was an investigation of the long-term reading performance of former Reading Recovery students in the Palm Beach School District. The review of related literature revealed that the Reading Recovery program was implemented as an intervention for seventh- and eighth-grade students in response to a need to address high numbers of students failing to learn to read (Pinnell, 1996). Irrespective of wide spread reservations in regard to its long-term benefits and cost efficacy, Reading Recovery continued to be widely implemented in schools across the United States. Information posted on the Reading Recovery Council National Association's Web site indicated that more than 2 million students were served in Reading Recovery programs nationwide during the 2006-2007 school year.

Educators familiar with Reading Recovery sites around the country studied the short-term benefits of Reading Recovery since its initial implementation in the United States in 1984 (Pinnell, 1996). This study tracked Reading Recovery students through the end of the 2006-2007 school year. Exploring the benefits of one early intervention reading program may contribute to gaining important insight to improve opportunities for all students to learn to read.

Similar to the current study, the available literature provided mixed evidence in support of claims of initial reading gains from Reading Recovery instruction (Browne, 1997) and indicated a need for further examination of the sustained effects of Reading Recovery instruction. Much of the early literature on the effects of Reading Recovery instruction concentrated more on first-grade pre- and posttest results and justifying the cost-effectiveness of Reading Recovery as an early intervention program. The review of literature for this study suggested the need for a more systemic examination of the long-

term effects of Reading Recovery instruction. The studies reviewed varied greatly in research design and methodology, including the number of years students were tracked, the measurements of reading performance, and statistical tests applied.

The FCAT was selected as the initial measurement tool for this study because it was the division-wide assessment instrument in 2007. It was the result of this test that identified those students in need of early intervention. Later, as the students reached seventh grade, it was the FCAT that was evaluated to determine student achievement and to evaluate curriculum. Did it effectively measure achievement for remedial reading students? A comprehensive picture of reading achievement is difficult to capture in any 55-item test.

The total reading section featured word analysis and comprehension. Those two components of the reading process are often considered part of a more traditional approach to reading, but they demand that a student demonstrate a grasp of skills such as phonemic awareness, letter-sound correspondence, syllabication, and syntax in order to read narrative and expository-graded passages fluently and accurately. Multiple-choice questions targeted a student's ability to distinguish main idea and detail; understand multiple meanings of vocabulary; and identify inference, sequence, and cause and effect to assess comprehension. In as much as reading is not a string of isolated skills, the FCAT measured students' ability in a particular time and place to integrate reading strategies to discover meaning. Perhaps, an instrument specifically designed to test multiple facets of reading in depth would have produced a totally different distribution of scores. A study that is descriptive by definition must describe what occurs as usual practice in a given place, not prescribe what occurs. Thus, the resulting FCAT scores allowed the reading supervisor to ascertain the impact of the instructional programs in

reading by scrutinizing the responses to test items, and the standardized instrument easily satisfied the test coordinator interested in comparing scores within the school division and the nation.

The research question for the study, What will be the effectiveness of a Reading Recovery program on improving the FCAT score of at-risk students?, helped summarize the results and conclusions. The results showed that the objectives of the Reading Recovery program matched the goals of the school. The results of the matching goals and objectives implied that the program would align to the need of the group. The evidence of the interviews and observation showed that the Reading Recovery program was successfully meeting program reading objectives. The Reading Recovery program improved student achievement in reading and comprehension levels. The researcher specifically noted that implementation of the reading program significantly reduced the number of at-risk students in Grades 7 and 8 of the identified middle school.

Research showed that the following four components are essential to have in an effective reading program: (a) professional development for continual learning, (b) curriculum and instruction, (c) assessment leading learning, and (d) the learning environment. Professional development provides teachers with the opportunity to learn how to provide comprehensive instruction for students that is research based. If teachers are provided with professional training in student instruction, they can help students become successful learners.

### *Discussion*

Unlike other Reading Recovery studies involving only successful students, the population under analysis was inclusive. Students who met the central tendencies of this research received Reading Recovery instruction in the county several years ago and had

taken the Stanford 9 test in the spring. Within the population were children who were successfully discontinued from the program and those who were not, children who needed support from Chapter I and those needing no support, and children who absorbed content in minimal time periods and those requiring time extensions. The inclusion of the entire spectrum of students as subjects permitted a more confident interpretation of results.

The mean of other Reading Recovery students, discontinued with more than 60 lessons, hovered around the median of the distribution of mean. Students discontinued with fewer or more than 60 lessons but with 1 to 2 years of Chapter I achieved means in the top half of the distribution.

Although two Reading Recovery groups who were not successfully discontinued had means slightly above the median, five groups not discontinued with any number of Lessons, despite 1 to 3 years of Chapter I, had means in the bottom half of the distribution, including the two lowest reading scores. The results from this study suggested that the achievement of the Reading Recovery population was solidly sustained for the long-term when students were successfully discontinued in any number of lessons and Reading Recovery instruction was buoyed with 1 to 2 years of Chapter I.

Because Reading Recovery prescribes a model of reading behaviors that can rapidly propel the lowest achieving eight graders toward literacy, the question arises as to why successful students might need subsequent support. Keeping in mind that the word *successful* might be a tenuous term for a group ranked the lowest, several elements of the study could be contributing factors.

Successfully discontinued from Reading Recovery indicated that a child understood and could apply reading strategies with a degree of independence comparable

to the average students in their homeroom class. But, new skills need practice. Momentum could be lost if the child returned to a classroom where Reading Recovery strategies were not reinforced either philosophically or practically by the classroom teacher. The results of this study within the context of the research question frames a discussion as it related to the implications of the short- and long-term reading performance of former Reading Recovery students.

The study participants came from the population of all seventh- and eighth-grade students eligible for supplemental reading support in Palm Beach County. Students who received Reading Recovery lessons represented the lowest scoring children in the bottom 20% of seventh- and eighth-grade classrooms. The wait-listed students represented the highest scoring children in the bottom 25%. A review of the FCAT reading scores revealed that those students who were former Reading Recovery students did not score significantly higher than those who were eligible for Reading Recovery but were wait listed. The FCAT scores of former Reading Recovery students reflected a slightly higher reading percentage overall. The Reading Recovery scores for the graduates of the Reading Recovery program several years prior were at 52% on the FCAT reading, vocabulary, and reading composite section. Meanwhile, the reading scores for those wait listed stood at 31% on the FCAT reading, vocabulary, and reading composite section.

The data from the tables offered an opportunity to merge recent knowledge about cognition, instruction, and culture with Reading Recovery to develop instructional systems that can make significant changes in the way the teaching and learning process for at-risk students is thought about and carried out in the schools. A key characteristic of the Reading Recovery program is its accelerative student achievement. There was a statistically significant difference in students' end-of-year reading percentages in the fall

for students who participated in Reading Recovery and those who were wait listed in the program.

### *Implications for Practice*

The results of this study presented intriguing possibilities for school personnel interested in or charged with the task of maximizing use of school time, talent, and resources. The two groups of students with the highest means successfully completed the Reading Recovery program in fewer than 61 lessons. The children or their Reading Recovery teachers had characteristics in common that deserve further evaluation and possible replication that enabled them to score well and within the ideal time frame. The scores of wait listed students correlated with scores in the top half of the distribution. Students should not begin a Reading Recovery program that cannot be completed before the close of the school year because the results showed that students with any number of lessons but not wait listed repeatedly produced disappointing scores at the bottom of the distribution. This suggested that the majority of students eligible for Chapter I in the primary grades may just not be developmentally ready to benefit from a 1-year concentrated boost to classroom instruction. Wait-listed Reading Recovery students also produced scores on a level of those who had received actual Reading Recovery instruction.

It would be advantageous for Chapter I teachers to package their programs for delivery consciously in 1- or 2-year time spans. This approach could enable Chapter I teachers to help transition Reading Recovery students as a matter of course as well as accommodate English as a second language or new students at any grade level. Such a role for Chapter I would utilize its strengths alone or as a valuable adjunct to Reading Recovery but not dilute its effectiveness by habitually assigning the same roster of low

achievers to the program. A plan such as this has implications for elementary classrooms.

In the middle school classroom, administrators suggested that curriculum supervisors support a slight realignment of the language arts and reading curriculum at the building level as recommended by Allington (2000). Administrators also suggested teachers at all grade levels create literacy-rich environments by actively engaging students in reading and writing tasks across the curriculum as implied by Barr (2000). Devoting blocks of time daily to the reading and discussion of quality literature, as advocated by Allington and McGill-Franzen (2000), identifying the instructional reading level of each student at the start of the school year, differentiating instruction, and supplying classroom libraries as well as school libraries with high-interest books at various independent reading levels are just a few of the concrete steps teachers may take to help students see themselves as readers. It would appear prudent to review the circumstances of students who scored in the lowest percentile rank in hopes of discerning a more effective means to guide subsequent student learning. Standardizing the screening devices used in the initial assessment of seven graders or refining the selection process for Reading Recovery and Chapter I students would help verify the identification of the lowest readers at the outset.

It is important to restate here that 42% of all students who participated in supplemental reading programs in Grades 7 and 8 were unable to participate in the follow-up study due to student and family mobility. Such an attrition rate weakens attempts by any school system to help its struggling students. Because instructional opportunities can frequently be lost during the readmittance process to a new school, moving during the early elementary years may compound the difficulties of low-achieving readers. A sense of community among division reading supervisors could be

fostered so, at least within a region or a state, a sense of shared responsibility for the reading success of all children could manifest particularly in caring attention to detail when placing low achievers who have moved.

#### *Recommendations for Further Research*

Which reading intervention works best for students at risk? Issues in reading worthy of research abound. These particular recommendations for further research as well as for school systems that regard the Reading Recovery program as a strong component of their instructional program in reading have emerged from the results obtained in this applied dissertation. The present study was a longitudinal study that incorporated descriptive data. The same data examined the IQ of participants as an alternative measure to provide useful information for primary-level curriculum coordinators among other school personnel.

Because it is reasonable to assume that most of the students under Reading Recovery instruction still attend school in Palm Beach County, the long-term could be expanded into a study that would track the reading achievement of Reading Recovery students in Grades 7, 8, and 9 or as frequently as the FCAT is administered.

Growth in reading achievement is a common focus for many classroom teachers as well as for remedial educators, but, as a tool to unlock meaning, reading can be examined in relationship to other content areas. Research examining the short- and long-term effects of reading achievement on math, social studies, and science could help administrators select appropriate tests and become more knowledgeable when interpreting test results. As Reading Recovery has expanded throughout the United States, elements of the program such as systematic observation of reading behaviors and an abundance of leveled texts in classrooms have become assimilated into classrooms as

general groups of best practice. Theoretically, struggling readers should thrive in classes with teachers who have strong reading backgrounds, promote literacy, and possess empathy for low achievers.

Research describing the effect of placing Reading Recovery students in literacy-rich, seventh- and eighth-grade classrooms could yield valuable information regarding the efficacy of grouping practices on students' sustained reading achievement. If it is true that 1 or 2 years of Chapter I maximize the progress of students in the lowest 25% of Grades 7 and 8, it would be most beneficial to discover the most opportune time to begin supplemental instruction. It is recommended that research be continue to determine what methods are in place in schools to monitor the progress of the lowest achieving students regularly.

Because education is a collaborative effort among many, is there a difference in the reading achievement of children in schools whose principal; guidance counselor; speech, reading, and special education specialists; physical therapist; social worker; and psychologist meet frequently to review the growth of individual students who struggle? As it relates to the long-term reading performance of the subjects in this study, the findings revealed persistent disparities across instruments and grade levels and across gender, ethnicity, and socioeconomic status lines. Although such disparities were not the major focus of this study, this was by far the most significant finding and warrants further investigation of the nature and causes of disparities among student groups.

The question Palm Beach County asked itself nearly 10 years ago was How can we accelerate the reading progress of our lowest-achieving children most efficiently? In response to that request, a new program, Reading Recovery, took its place in this language arts curriculum next to the old standby, Chapter I. In the 4 years of Reading

Recovery's implementation, students took instruction at all Palm Beach County middle schools. Four years later, an evaluation of test results revealed that the two programs, especially in concert, produced slight differences in scores.

### *Conclusion*

This study represented the school's first attempt to evaluate the effectiveness of a reading program that was implemented by every teacher in Grades 7 and 8. The evaluation results suggested that the Reading Recovery program was successful in reducing the number of at-risk students. The findings also suggested that the teachers who participated in the program viewed the program as beneficial. The findings further indicated that the teachers now view assessments as an important piece in identifying individual student needs. Evaluation of the literacy environment in the classrooms indicated that teachers who were using small-group instruction were focused on meeting the needs of individual students. Because there were a large number of students who did achieve academic proficiency, the researcher concluded that the Reading Recovery program did make statistically significant differences in improving the reading of at-risk students.

## References

- Adams, M. (1995). *Beginning to read: Thinking and learning about print*. Cambridge, MA: The MIT Press.
- Aldridge, J. (2000). *Current issues and trends in education* (2nd ed.). Boston: Allyn & Bacon.
- Allington, R. L. (2000). The reading instruction provided readers of different reading abilities. *Elementary School Journal*, 83, 548-559.
- Allington, R. L., & McGill-Frazen, A. (2000). School response to reading failure: Instruction for Chapter I and special ed students in Grades 2, 4, and 8. *Elementary School Journal*, 89, 529-542.
- Allington, R. L., Stuetzel, H., Shake, M., & Lamarche, S. (2000). What is remedial reading? A descriptive study. *Reading Research and Instruction*, 26, 15-30.
- Allington, R. L., & McGill-Franzen, A. (2000). Placing children at risk: Schools respond to reading problems. In R. Donmoyer & R. Kos (Eds.), *At-risk students: Portraits, policies, programs, and practices* (pp. 197-217). Albany, NY: SUNY Press.
- Allington, R. L., & Walmsley, S. A. (1995). No quick fix: Where do we go from here? In R. L. Allington & S. A. Walmsley (Eds.), *No quick Fix: Rethinking literacy programs in American's elementary schools* (pp. 253-264). New York: Teachers College Press.
- Askew, B. J., Fountas, I. C., Lyons, C. A., Pinnell, G. S., & Schmitt, M. C. (1998). *Reading Recovery review: Understanding outcomes and implications*. Columbus, OH: Reading Recovery Council of North America.
- Barnes, B. (1997). But teacher you went right on: A perspective on Reading Recovery. *The Reading Teacher*, 50, 284-292
- Barr, R. (2000). Classroom reading instruction from a sociological perspective. *Journal of Reading Behavior*, 14, 375-389.
- Bean, R. M., Eichelberger, R. T., Lazar, M. K., & Zigmond, N. (1995). In class or pullout: Effects of setting on the remedial reading program. *Journal of Reading Behavior*, 23, 445-464.
- Beech, J. R., & Harding, L. M. (1995). Phonemic processing and the poor reader from a developmental lag viewpoint. *Reading Research Quarterly*, 19, 357-366.
- Browne, A. (1997). Teaching and learning in Reading Recovery: Response to But teacher you went right on. *The Reading Teacher*, 50, 294-303.

- Carter, L. (1996). The sustaining effects of compensatory and elements education. *Educational Researcher, 13*, 13-14.
- Center, Y. (1995) An evaluation of Reading Recovery. *Reading Research Quarterly, 30*, 240-263.
- Chall, J. S. (1996). Learning to read: The great debate. In H. E. Mitzel, J. H. Best, & W. Rabinowitz (Eds.), *Encyclopedia of educational research* (p. 1538). New York: The Free Press.
- Chard, D. J., Kame'enui, E. J., & Simmons, D. C. (1994). Translating research into basal reading programs: Application of curriculum design. *LD Forum, 20*(1), 9-13.
- Clay, M. M. (1995a). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.
- Clay, M. M. (1995b) *Observing young readers: Selected papers*. Portsmouth, NH: Heinemann.
- Clay, M. M. (1996a). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.
- Clay, M. M. (1996b). *Reading Recovery: A guidebook for teachers in training*. Portsmouth, NH: Heinemann.
- Clay, M. M. (1998). *The early detection of reading difficulties*. Portsmouth, NH: Heinemann.
- Conoley, J. C., & Kramer, J. J. (Eds.) (1998). *The tenth mental measurements yearbook*. Lincoln: University of Nebraska-Lincoln, Buros Institutes of Mental Measurements.
- Deegan, D. H. (1995). The necessity of debate: A comment on commentaries. *The Reading Teacher, 48*, 688-695.
- Deford, D. E. (1996). Validating the construct of theoretical orientation in reading instruction. *Reading Research Quarterly, 20*, 351-367.
- Donmoyer, R., & Kos, R. (Eds.). (1996). *At-risk students: Portraits, policies, programs, and practice*. Albany, NY: SUNY Press.
- Ehri, L. C. (1996). Development of spelling knowledge and its role in reading acquisition and reading disability. *Journal of Learning Disabilities, 22*, 356-365.
- Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational*

*Psychology*, 92, 605-619.

- Felton, R. H. (1995). Effects of instruction on the decoding skills of children with phonological-processing problems. *Journal of Learning Disabilities*, 26, 583-589.
- Fitzgerald, J. (1999). What is this thing called balance? *The Reading Teacher*, 53, 100-107.
- Florida Department of Education. (1996). *Sunshine state standards*. Tallahassee, FL: Author.
- Foorman, B. R., Francis, D. J., Novy, D. M., & Liberman, D. (1996). How letter-sound instruction mediates progress in first-grade reading and spelling. *Journal of Educational Psychology*, 83, 456-469.
- Foorman, B. R., & Liberman, D. (1996). Visual and phonological processing of words: A comparison of good and poor readers. *Journal of Learning Disabilities*, 22, 349-355.
- Goodman, K. S. (1998). Behind the eye: What happens in reading. In H. E. Mitzel, J. H. Best, & W. Rabinowitz (Eds.), *Encyclopedia of education research* (p. 1537-1539). New York: The Free Press.
- Goyen, J. (2000). Reading methods in Spain: The effect of a regular orthography. *The Reading Teacher*, 42, 370-373.
- Gredler, G. R. (1997). Intervention programs for youth children with learning problems. *Problems in the Schools*, 34, 161-169.
- Hiebert, E. H. (1995). Reading Recovery in the United States: What differences does it make to an age cohort? *Educational Researcher*, 25(7), 26-28.
- Hobsbaum, A. (1996). Scaffolding in Reading Recovery. *Oxford Review of Education*, 22(1), 17-35.
- Hurford, D. P., Schauf, J. D., Bunce, L., Blaich, T., & Moore, K. (1995). Early identification of children at risk for reading disabilities. *Journal of Learning Disabilities*, 27, 371-382.
- Iversen, S., & Tunmer, W. E. (1996). Phonological processing skills and the Reading Recovery program. *Journal of Educational Psychology*, 85, 112-126
- Juel, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly*, 31, 268-289.
- Knight, C. C., & Fischer, K. W. (1995). Learning to read words: Individual differences in developmental sequences. *Journal of Applied Developmental Psychology*, 13,

377-404.

- Maclean, R. (1997). Two paradoxes of phonics. *The Reading Teacher*, 42, 514-517.
- McGuiness, D., McGuiness, C., & Donohue, J. (1995). Phonological training and the alphabet principle: Evidence for reciprocal causality. *Reading Research Quarterly*, 30, 830-852.
- McKenna, M. C., Stahl, S. A., & Reinking, D. (1995). A critical commentary on research, politics, and whole language. *Journal of Reading Behavior*, 26, 211-233.
- Ohio State University. (1992). *Reading Recovery executive summary*. Columbus, OH: Author.
- Perfetti, C., & Curtis, M. (1995). Development of children's reading self-concepts and their relationship with reading achievement. *Journal of Educational Psychology*, 87, 155-169.
- Pikulski, J. (1995). Preventing reading failure: Review of five effective programs. *The Reading Teacher*, 48, 30-39.
- Pinnell, G. S. (1996). Reading Recovery: Helping at-risk children learn to read. *Elementary School Journal*, 90, 161-183.
- Pinnell, G. S. (2005). *Reading Recovery: A review of research* (Educational Report No. 23). Columbus, OH: Martha L. King Language and Literacy Center.
- Pinnell, G. S., Lyons, C. A., Deford, D. E., Bryk, A. S., & Seltzer, M. (1995). Comparing instructional models for the literacy education of high-risk ninth graders. *Reading Research Quarterly*, 29, 9-39.
- Pressley, M. (1995). State-of-the-science primary grades reading instruction of whole language. *Educational Psychologist*, 29, 211-215.
- Rasinski, T. V. (1995). Commentary on the effects of Reading Recovery: A response to Pinnell, Lyons, Deford, Bryk, and Seltzer. *Reading Research Quarterly*, 30, 276-277.
- Reading Recovery Council of North America. (2001). *Reading Recovery*. Retrieved May 13, 2007, from <http://www.readingrecovery.org/sections/reading/index.asp>
- Reutzel, D. R. (1999). On balanced reading. *The Reading Teacher*, 52, 322-324.
- Ross, S. M. (1995). Increasing the academic success of disadvantaged children: An examination of alternative early intervention programs. *American Educational Research Journal*, 32, 773-800.

- Rumelhart, O. E. (1996). Toward an interactive model of reading In. H. E. Mitzel, J. H. Best, & W. Rabinowitz (Eds.), *Encyclopedia of educational research* (p. 1537). New York: The Free Press.
- Santa, C. M., & Hoiem, T. (1999). An assessment of early steps: Program for early intervention of reading problems. *Reading Research Quarterly, 34*, 54-79.
- Scarborough, J. S., Dobrich, W., & Hager, M. (1996). Preschool literacy experience and later reading achievement. *Journal of Learning Disabilities, 24*, 508-511.
- Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly, 30*, 958-996.
- Silvaroli, N. J. (1997). *Classroom reading inventory*. Dubuque, IA: Times Mirror Higher ED Group.
- Simmons, D. C., & Kame'enui, E. J. (Eds.). (1998). *What reading research tells us about children with diverse learning needs*. Mahwah, NJ: Erlbaum.
- Singer, J., & Ruddell, R. (Eds.). (1997). Theoretical models and processes of reading. In J. E. Mitzel, J. H. Best, & W. Rabinowitz (Eds.), *Encyclopedia of educational research* (p. 1536). New York: The Free Press.
- Slavin, R. E. (1995). Making Chapter I make a difference. *Phi Delta Kappan, 72*, 586-589.
- Slavin, R. E. (2005). *Effective classroom programs for students at risk*. Boston: Allyn & Bacon.
- Spiegel, D. L. (1995). A comparison of traditional remedial programs and Reading Recovery: Guidelines for success for all programs. *The Reading Teacher, 49*, 86-96.
- Spiegel, D. L. (1998). Response to Deegan. *The Reading Teacher, 48*, 696-698.
- Stahl, S. A., McKenna, M. C., & Pagnucco, J. R. (1995). The effects of whole language instruction: An update and a reappraisal. *Educational Psychologist, 29*, 175-185.
- Stanovich, K. E., & Siegel, L. S. (1996). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable differences model. *Journal of Educational Psychology, 86*, 24-53.
- Thistlethwaite, L., & Mason, M. (1995). Chapter I reading programs: Do student/family factors make a difference? *Reading Horizons, 33*, 258-277.

- Torgensen, J. (1998). Catch them before they fall: Identification and assessment to prevent reading failure in young children. *American Educator*, 22, 32-39.
- U.S. Department of Education. (2005). *No child left behind*. Retrieved June 22, 2005, from <http://www.ed.gov/nclb/landingjhtml>
- Vandervelden, M. C., & Siegel, L. S. (1995). Phonological recoding and phoneme awareness in early literacy: A developmental approach. *Reading Research Quarterly*, 30, 854-875.
- VanKleeck, A. (1998). Preliteracy domains and stages: Laying the foundations for beginning reading. *Journal of Children's Communication Development*, 20(1), 33-51.
- Worhan, S. C. (1995). *Measurement and evaluation in early childhood education*. Englewood Cliff, NJ: Prentice-Hall.