

# Collective Teacher Efficacy And Reading Achievement For Hispanic Students In Reading First And Non-reading First Schools In Southwest Florida

2005

Larry Hylemon  
*University of Central Florida*

Find similar works at: <http://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

 Part of the [Curriculum and Instruction Commons](#)

## STARS Citation

Hylemon, Larry, "Collective Teacher Efficacy And Reading Achievement For Hispanic Students In Reading First And Non-reading First Schools In Southwest Florida" (2005). *Electronic Theses and Dissertations*. 571.  
<http://stars.library.ucf.edu/etd/571>

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of STARS. For more information, please contact [lee.dotson@ucf.edu](mailto:lee.dotson@ucf.edu).

COLLECTIVE TEACHER EFFICACY AND READING ACHIEVEMENT  
FOR HISPANIC STUDENTS IN *READING FIRST* AND *NON-READING FIRST*  
SCHOOLS IN SOUTHWEST FLORIDA

By

LARRY VAN HYLEMON  
B.A. East Carolina University, 1971  
M.A. East Carolina University, 1976  
C.A.S. East Carolina University, 1979

A dissertation submitted in partial completion of the requirements  
for the degree of Doctor of Education  
in the Department of Educational Studies  
in the College of Education  
at the University of Central Florida  
Orlando, Florida

Fall Term  
2005

Major Professor: Larry C. Holt

©2005 Larry Van Hylemon

## ABSTRACT

This study examined explicit and systematic reading instruction for Hispanic, limited English, lower SES students in a southwest Florida school district. Additionally, collective teacher efficacy was assessed to determine if differences existed between a Reading First and a non-Reading First school. A total of 68 students participated in the study and were divided equally between a Reading First and non-Reading First school. While the Reading First school concentrated resources on grades 2 and 3, all students received some degree of Reading First strategies. Available reading measures for analyses at the time of the study included the Group Reading Assessment and Diagnostic Evaluation (GRADE), the Florida Comprehensive Assessment Test (FCAT), and the Dynamic Indicators of Basic Early Literacy Skills (DIBELS).

Results from two-way analyses of variance (ANOVA) did not reveal significant differences between the two groups on a pre and posttest administration of the GRADE. When the two groups were combined, repeated measures ANOVA indicated a significant difference between pre and post administrations of the GRADE with the posttest being significantly higher. Grades 4 and 5 showed significant gains between pre and posttest while grades 2 and 3 did not. It was hypothesized that as students' English proficiency improved, reading instruction became more meaningful and thus positively influenced the posttest. Both the GRADE and the DIBELS Oral Reading Fluency test were significantly correlated with the FCAT. Regression analyses revealed that both were significant predictors individually and combined of the FCAT reading score.

Collective teacher efficacy was assessed with the short-form of the Collective Teacher Efficacy Scale (CTES). A total of 38 Reading First teachers and 30 non-Reading

First teachers completed the CTES and a brief biographical questionnaire. Results indicated significantly higher collective teacher efficacy in the non-Reading First school. Findings did not support the hypothesis that the intensity of the Reading First program and the teacher training required would result in significantly higher collective efficacy. Differences in level of teacher education, experience, and years teaching in a school were speculated as potential variables influencing the level of collective efficacy. Implications from this study were discussed along with recommendations for future research.

To Deborah, Chris, and Lauren for their love, patience, and constant encouragement.  
And, to the memory of Dr. Betty Jane Corwin who saw potential where few others chose  
to look.

## ACKNOWLEDGMENTS

This study was completed with the assistance and input from many individuals who made contributions throughout its completion. Dr. Larry Holt served as committee chair and provided invaluable help during the entire process. Further, he did so with utmost professionalism and with remarkable patience. I am particularly indebted to him for all his efforts and the privilege of having him on my committee. Dr. George Pawlas was continuously willing to offer any help possible and was extremely astute in identifying needed changes during his readings of the document. I am very thankful for his promptness and insightful recommendations. Dr. Lea Witta was first my statistics professor who created an air of excitement for a subject that previously had been dreaded. As a dissertation committee member, she provided outstanding expertise in the analyses and interpretation of data. Dr. Maria Torres brought with her an extensive level of knowledge and experience in the area of second language learning. Her comments and suggestions reflected her intense caring about the education of second language students and were extremely valuable in improving the quality of this study. I am most grateful for her willingness to serve on my committee.

Many individuals in the Collier County school district either directly or indirectly contributed to this study. I am particularly indebted to Ms. Nancy Smith whose original ideas formed the basis for this study. Mrs. Beth Thompson patiently answered innumerable questions about the reading instruction in Collier County and the Reading First program. I am very grateful to have had her input. To Dr. Connie Helton and Dr. Jan Messer and their school staffs, I am most appreciative of their willingness to participate in this study and their efforts in procuring the necessary data from their schools. In

addition, I am thankful for the support of the district Research Oversight Committee who provided final approval for this study and particularly, Dr. Vivian Shayne. She kindly reviewed the data analyses and offered insights and suggestions that were extremely valuable in further clarifying the implications of this study.

Finally, to the Golden Gate Elementary Multidisciplinary Team members who were often a captive audience and to my Cohort Four friends, Marybeth, Nathan, and Diana, I wish to acknowledge my sincere appreciation to all for your encouragement and friendship throughout the course of study and this dissertation.

## TABLE OF CONTENTS

LIST OF TABLES .....	xi
LIST OF ABBREVIATIONS.....	xii
CHAPTER ONE PROBLEM STATEMENT AND DESIGN COMPONENTS .....	1
Introduction.....	1
Statement of Purpose .....	4
Research Questions.....	4
Definition of Terms.....	5
Assumptions and Limitations .....	7
Data Collection and Instrumentation .....	8
Data Analysis.....	11
Organization of the Study .....	12
CHAPTER TWO LITERATURE REVIEW .....	13
Introduction.....	13
Theoretical Underpinnings of Second Language Literacy .....	13
Educational Program Models.....	17
Systematic and Explicit Reading Instruction.....	23
The National Reading Panel’s Findings .....	24
Phonemic Awareness.....	25
Phonics.....	26
Fluency.....	28
Comprehension .....	30

Vocabulary .....	30
Text Comprehension.....	31
Explicit Reading Instruction and the ESL Student .....	32
Phonemic Awareness and Phonics.....	33
Fluency.....	36
Vocabulary .....	37
Text Comprehension.....	39
Conclusions.....	41
Collective Teacher Efficacy.....	42
Measurement of Collective Teacher Efficacy.....	45
CHAPTER THREE METHODOLOGY .....	47
Research Design.....	47
Participating Schools .....	47
Participants.....	49
Procedures.....	54
Instrumentation .....	54
Data Collection .....	59
Variables .....	60
Data Analysis .....	61
CHAPTER FOUR ANALYSIS OF DATA.....	63
Results as Related to Research Questions .....	63
Question One .....	64
Question Two.....	69

Question Three.....	73
Question Four.....	74
Question Five.....	76
Question Six.....	78
CHAPTER FIVE CONCLUSIONS, DISCUSSION, AND RECOMMENDATIONS ...	83
Summary of Findings.....	83
Discussion.....	87
Reading First and ESL Students.....	87
Collective Teacher Efficacy.....	90
Conclusions.....	92
Recommendations for Future Research.....	93
APPENDIX A COLLECTIVE TEACHER EFFICACY SCALE-SHORT FORM.....	97
APPENDIX B TEACHER QUESTIONNAIRE.....	100
APPENDIX C PRINCIPAL PERMISSION FORM – READING FIRST SCHOOL ...	102
APPENDIX D PRINCIPAL PERMISSION FORM – NONREADING FIRST SCHOOL .....	105
APPENDIX E TEACHER CONSENT FORM.....	108
APPENDIX F IRB APPROVAL LETTER.....	110
LIST OF REFERENCES.....	112

## LIST OF TABLES

Table 1 Demographics of Elementary A and Elementary B Schools.....	49
Table 2 Student Demographics.....	51
Table 3 Teacher Demographics .....	53
Table 4 Developmental Reading Assessment Levels from the 2002-03 Academic Year	55
Table 5 Elementary A and B Descriptive Statistics for Total Grade Standard Scores.....	66
Table 6 Mean Number of Items and Percentages Answered Correctly on GRADE Subtests .....	68
Table 7 Two-Way ANOVA for GRADE Pre and Posttest Scores, Grades 2-5, Group, and Interaction Effects.....	72
Table 8 Repeated Measures ANOVA.....	73
Table 9 DIBELS, FCAT, and GRADE Correlations.....	76
Table 10 Predictors of FCAT Developmental Scores for Reading First and Non-Reading First Groups .....	78
Table 11 Item Statistics for the Reading First CTES.....	80
Table 12 Item Statistics for the Non-Reading First CTES .....	81
Table 13 Item Statistics for Both the Reading First and Non-Reading First CTES .....	82

## LIST OF ABBREVIATIONS

CTE: Collective teacher efficacy

ESL: English as a second language

ESOL: English for speakers of other languages

L1: First or heritage language

L2: Second language

NRP: National Reading NRP

SES: Socioeconomic status

## CHAPTER ONE

### PROBLEM STATEMENT AND DESIGN COMPONENTS

#### Introduction

The number of English language learners in the U.S. public schools has increased dramatically over the last decade. During the 2000-2001 school year, there were more than 4.5 million students learning English. This number represented an increase of 32 percent over the 1997-98 school year (Miller, 2003). The 2002 National Center for Education Statistics (NCES) report revealed that 9.64% of public school students were acquiring English as a second language (ESL) and among those, 79% were Spanish speakers. The rise in enrollment for these students has been continuous and particularly in the state of Florida. During the 2001-2002 school year, 20.4% of Florida's public school students were identified as Hispanic (NCES, 2002). As the percentage of ESL students continues to grow in U.S. public schools, the need for effective and appropriate instructional practices becomes increasingly significant. As noted by Miller, this issue has been further complicated by state budget decreases and new legislative requirements.

The No Child Left Behind (NCLB) (U.S. Department of Education, 2004) legislation of 2001 mandated that all students, including those acquiring English, demonstrate significant academic gains. The law specifically required that all students show annual yearly progress and read on grade level by the 2013-2014 academic year. This requirement presents a particular challenge to a group of students who have shown

increasingly wider gaps in achievement when compared to native English speakers (Lopez & Tashakkori, 2003). ESL students are confronted with the task of learning English while acquiring academic skills simultaneously. According to Collier (1989), young children entering the U.S. with little or no instruction in their primary language may need 7 to 10 years to attain average reading achievement in English. These students face considerable pressure to acquire reading skills at a rapid pace with the expectation that they meet the same academic standards as monolingual students (Miller, 2003). According to a National Assessment of Educational Progress report (NAEP) (U.S. Department of Education, 1999), approximately 60% of Hispanic students at the fourth grade level were below basic proficiency in reading. In Florida, third grade students who do not pass the Florida Comprehensive Assessment Test (FCAT) risk multiple grade retentions. Hispanic students accounted for 29% of the retained third graders in Florida for the 2003-04 academic year (FASP, 2003).

Historically, there have been numerous, and sometimes conflicting programmatic and instructional approaches implemented to address ESL students' English literacy skills. Programs have ranged from English only to first language only instruction with variations in the amount of time instruction was provided in the first language. More recently and at the national level, the Reading First program was initiated in all states as a part of the NCLB act with federal funds to support it. The program's research-based strategies provide systematic and explicit instruction in the areas of phonemic awareness, phonics, fluency, vocabulary, and comprehension (Just Read, Florida!, n/d). As implemented in Florida, Reading First was developed to increase reading skills for both monolingual and ESL students alike. Teachers have received specific training in reading

instruction that coincides with the theoretical foundations and research identified by Reading First. Due to its recency in implementation, little was known about the program's effectiveness with ESL students. In view of the reported acquisition time by Collier (1989) for some ESL students to master academic skills, together with the No Child Left Behind mandates, the need to critically examine reading methods and insure that the most effective are being taught to ESL students would seem more paramount now than ever.

In addition to instructional methodologies and program types, numerous other factors contribute to student success. One factor that has shown a positive relationship with student achievement is collective teacher efficacy (CTE). Based upon Bandura's (1977, 1986, & 1997) social cognitive theory and research on individual teacher efficacy, collective efficacy refers to the belief of teachers in a school that they will, as a whole, have a positive effect on student outcomes (Goddard, Hoy, & Hoy, 2000). According to Goddard et al., collective efficacy has received relatively little research when compared to a much larger body of studies examining the relationship between student achievement and teacher self-efficacy. Further, there were few studies examining the relationship of CTE to school-wide implemented programs such as Reading First. In Reading First schools, it would seem logical that teachers must, as a group, perceive that they have the abilities to implement the specific strategies if the program is to be successful. For those schools where the majority of teachers have been trained in Reading First, it could be questioned whether the systematic and highly explicit nature of Reading First instruction could foster collective efficacy beliefs.

### Statement of Purpose

The purpose of this study was twofold: First, it attempted to determine whether Reading First ESL participants showed different levels of reading achievement when compared to a group of ESL students who did not attend a Reading First school. Second, it compared teachers' collective efficacy between the Reading First and non-Reading First schools. The researcher anticipated that the outcome of this study could provide contributions to the current knowledge about the benefits of systematic and explicit reading instruction for ESL students. Additionally, findings could have implications pertaining to the relationship between the Reading First program's philosophy and methodology, and collective teacher efficacy beliefs.

### Research Questions

1. What are the reading achievement levels of a Reading First group and a non-Reading First group of ESL students after a posttest on the Total Test and the individual subtests of the Group Reading Assessment and Diagnostic Evaluation (GRADE), (Williams, 2002)?
2. Is there a significant difference between the two groups in the amount of progress made during the 2004-2005 school year as shown by the GRADE?
3. What is the correlation between scores on the Dynamic Indicators of Basic Early Literacy Skills (DIBELS), (Good & Kaminski, 2001) and GRADE scores for the Reading First group?

4. What is the correlation between Total Score on the GRADE and the Florida Comprehensive Assessment Test (FCAT) for third, fourth, and fifth grade students in the Reading First and non-Reading First groups?
5. Can the GRADE and/or DIBELS scores serve as predictors of the FCAT Reading score?
6. Is there a significant difference in collective teacher efficacy, as assessed by the Collective Teacher Efficacy Scale (CTES) (Goddard, 2002), between the Reading First and the non-Reading First school?

#### Definition of Terms

1. English as a Second Language (ESL) – For purposes of this study, the definition of ESL was identical to that of the Reading First program in defining limited English proficient students (Just Read Florida!, n/d). As such, ESL students are those who were not born in the U.S. and have a native language other than English, or were born in the U.S. but whose language at home is other than English. As a result of these factors, ESL students have difficulty speaking, understanding, reading, and/or writing English.

2. English for Speakers of Other Languages (ESOL) – The Florida Department of Education (2000-2001) has defined ESOL as “instruction in English/Language Arts, regardless of delivery model approach” (p.3). For this study, ESOL is used to identify programmatic and instructional services for ESL students as provided in the specific district where the study was implemented. Participants in ESOL have met state and local

eligibility requirements and receive varying levels of academic and language interventions. Services range from self-contained placement, individual and small group tutoring, to yearly monitoring of academic progress by ESOL personnel. Students in this study received inclusion ESOL services.

3. Reading First – As a part of the NCLB Act, Reading First is a research based combination of systematic and explicit strategies for reading instruction. In Florida, it has been implemented by Just Read, Florida! (n/d) through the Florida State Department of Education. Reading First is a K-3 program designed to prevent reading difficulties rather than remediate them. Instruction is provided in the regular classroom setting by trained regular classroom teachers. Major components emphasized in the program include phonological awareness, phonics, vocabulary, comprehension, and fluency. The program's content also addresses spelling and writing skills. Additionally, Reading First teachers are trained to provide specialized strategies for ESL students. Reading First does not require specific materials for reading instruction but does train teachers to use books and other materials adopted by the school and district.

4. Collective Teacher Efficacy (CTE) – According to Bandura (1997), collective teacher efficacy is defined as “a group's shared belief in its conjoint capabilities to organize and execute the courses of action required to produce given levels of attainments” (p. 477). Bandura further stated that collective efficacy is an evolving attribute of the group and not merely a summed total of the individuals' personal efficacy beliefs. For purposes of this study, and consistent with Goddard, Hoy, and, Hoy (2000) and Goddard's (2002) definition, collective efficacy was defined as the beliefs of teachers

in a school that their efforts as a collective whole would have a beneficial and positive effect on student outcomes.

### Assumptions and Limitations

The following assumptions and limitations appear applicable to this study:

1. The samples selected for this study were lower SES, elementary school age, Spanish-speaking students in Southwest Florida who had demonstrated limited English proficiency. Generalizations from the findings are limited to similar populations.
2. There is limited ability to insure that teachers will use Reading First strategies consistently throughout the school year. However, assessment of Reading First students with the DIBELS four times during the year serves as a monitoring process. Students who are consistently deficient in skills would alert the school's Reading First coach and school administration of possible instructional deficiencies.
3. There is limited ability to insure equal quality of ESOL tutoring and academic instruction within the individual classes.
4. Collective teacher efficacy, reading achievement, and Reading First methodology are not necessarily cause and effect relationships. Other factors may exist that were not identified in this study, such as student ability, the influence of previous teachers, parental involvement in the students' education, and the individual school's leadership characteristics.

5. The measurement of collective teacher efficacy is limited to that as defined and validated by Goddard, Hoy, and Hoy (2000), and Goddard (2002).
6. It is assumed that Reading First trained teachers conscientiously implemented the program's strategies to students selected for this study. It is also assumed that the DIBELS, FCAT and GRADE tests were administered in a standardized fashion and subsequently scored and recorded accurately in the district's database.
7. A percentage of students in this study had some formal schooling in countries other than the U.S. Academic records from these countries were not available for this study. As a result, the actual length of formal schooling for all students could not be determined and could vary according to grade levels. Establishing the equivalency of reading levels prior to the study would contribute to reducing this limitation.

#### Data Collection and Instrumentation

Permission was obtained from a southwest Florida school district's Research Oversight Committee to conduct this study. Principals from a Reading First (Elementary A) and a non-Reading First (Elementary B) school gave permission for their schools to participate and to release student identification numbers. Copies of the principal permission forms are provided in Appendices C and D. Student identification numbers were then used to access test scores and other data from the district's electronic databases.

At Elementary A, a sample of 34 students was identified with an equal number located at Elementary B who met criteria for selection. Subjects were selected who met the following: (a) were limited English proficient as defined by the Bilingual Verbal Ability Tests (BVAT) (Muñoz-Sandoval, Cummins, Alvarado, & Ruef, 1998), (b) participated in ESOL during the 2004-2005 school year, (c) had no prior grade retentions, (d) were not identified as learning disabled, language impaired, or emotionally handicapped (e) had been tested with the Group Reading Assessment and Diagnostic Evaluation (GRADE) (Williams, 2002) during May 2004 and again in May 2005, (e) were students in grades 2 through 5 in either Elementary A or B, (f) were determined to have Spanish as their first language, and (g) participated in the district's free or reduced cost lunch program. Students meeting all criteria but who had not attended the Reading First school for the entire year were excluded. Students in both groups were participants in regular mainstream classrooms and received services through the ESOL program during the 2004-2005 school year. In both schools, ESOL was an English only program and provided inclusion tutoring by paraprofessionals under the direction of the classroom teacher.

In addition to the GRADE and BVAT as described above, other instruments used in this study included the Developmental Reading Assessment (DRA), (Beaver, 1997), the Flynt-Cooter Reading Inventory (FCRI) (Flynt & Cooter, 1998), the Florida Comprehensive Assessment Test (FCAT), the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2001), and the Collective Teacher Efficacy Scale – Short Form (CTES) (Goddard, 2002). Additionally, a brief biographical Teacher Questionnaire (Appendix B) was developed by the researcher and given to all teachers at

both schools for completion. Data from the DRA, and FCRI were used to first establish reading levels for the two groups during the 2003-2004 school year and at the initial participation of students in the Reading First program. Test data from the GRADE administered in May 2004 and May 2005 were obtained from the school district's electronic database *Data Warehouse*. The two administrations of the GRADE served as pre- and posttest measures for gauging reading achievement.

The DIBELS assessment was an integral and state required part of the Reading First program. All 34 Elementary A participants were administered the DIBELS during May 2005. For students in grade 2, scores were available and analyzed on the DIBELS Nonsense Word Fluency and Oral Reading Fluency tests. For students in grades 3-5, only the Oral Reading Fluency test from the DIBELS was administered.

To assess collective teacher efficacy in each school, a packet for each teacher was created and delivered to both Elementary A and B. Each packet contained a Teacher Questionnaire (Appendix B), the collective teacher efficacy measure (CTES) (Appendix A), two Teacher Consent forms (Appendix E), and an envelope. All packets were delivered two weeks before the end of the teachers' contract. The teachers returned their sealed envelopes to each principal who then provided them to the researcher at the end of the final week of school in May 2005. Teachers at Elementary A and B at all six grade levels and the Exceptional Student Education programs completed the CTES and Teacher Questionnaire except for third grade teachers at Elementary B. By the end of May 2005, no CTES or questionnaires had been returned by this group of teachers.

## Data Analysis

All data were entered into a research database using the *Statistical Package for the Social Sciences: Graduate Pack 11.5 for Windows* (SPSS, 2002) for the purpose of analysis. Data were grouped according to the questions of this study. Significance for all statistical analyses was determined at the .05 level.

The first research question of the study addressed the reading achievement levels of the two groups of students after the final administration of the GRADE in May 2005. Descriptive statistics including means and standard deviations were calculated for each group on the Total GRADE test and the individual subtests.

The second research question asked if a significant difference was seen between the two groups in the amount of reading progress made during the 2004-05 school year. To address this question, two-way analyses of variance (ANOVA) and repeated measures ANOVA were calculated on the basis of GRADE scores from the 2004 and 2005 administrations. Between-subjects factors included group participation, grade levels, and comparison of grade level combinations including 2-3 and 4-5.

For the third research question, it was asked what the correlation was between the DIBELS scores and the GRADE scores for the Reading First group. Pearson product-moment correlation was used to address this question.

The fourth research question also pertained to correlation and specifically asked if there was a relationship between the Total Score on the GRADE and the FCAT scores for the third, fourth, and fifth grade students in both groups. Again, Pearson product-moment correlation was used to address this question.

For research question five, the researcher asked if the GRADE and DIBELS scores could serve as predictors of the FCAT scores for students in grades 3-5. Linear regression analyses were performed for the combinations of GRADE/FCAT and DIBELS/FCAT scores. Multiple regression was done to address the predictability of the combined DIBELS and GRADE scores for FCAT scores.

The final research question asked if there was a significant difference in collective teacher efficacy between the Reading First school and the non-Reading First school based upon the CTES. The teachers in each school were considered as two independent groups and to address this question, a t-test for independent means was calculated.

#### Organization of the Study

The organization of this study was causal-comparative and employed a nonequivalent (pretest and posttest) control group design (Creswell, 2003). This design appeared most appropriate for addressing questions about the statistical significance of differences in reading achievement between two naturally occurring groups that could not be randomly assigned to experimental and control groups. The southwest Florida district designated all schools receiving Title I funds as Reading First schools. As a result, the researcher was not able to randomly place students in groups. Subjects for the two groups were matched in an effort to limit as much as possible the influence of extraneous variables including level of English proficiency, the impact of prior grade retentions, literacy in the students' heritage language, socioeconomic status, participation in an ESOL program, and participation in special educational programming.

## CHAPTER TWO

### LITERATURE REVIEW

#### Introduction

In an effort to establish the relevance of this study and to gauge its potential significance, particular attention was given to research conducted during the past two decades. Search terms including bilingual education, second language learners, limited English proficient, immersion, teacher efficacy, and collective teacher efficacy were used. Terms were then refined to include specific aspects of reading instruction for ESL students. Emerging from this search was evidence of a sizable body of research directed at the effects of program types where instruction in the students' native language (L1) was varied and then related to reading achievement. Some basic theoretical models for second language acquisition and literacy were also identified. As a second aspect of this review, efforts were made to identify studies examining the effectiveness of systematic and explicit reading instruction for Spanish-speaking ESL students. Finally, attention was given to collective teacher efficacy with emphasis on its relationship to student achievement and teaching practices. Results and implications from these avenues of investigation are addressed individually.

#### Theoretical Underpinnings of Second Language Literacy

Constantino (1999) provided a comprehensive review of the literature pertaining to reading for second language (L2) learners and included a critical discussion of the

major theories believed to underlie the acquisition of a second language. Her first premise was that the primary prerequisite for reading was the knowledge of a language. In acquiring language, she first identified “foundational theories” that included researchers such as Chomsky (1959), Lenneberg (1967), and Ervin-Tripp (1973). Based upon a review of these theorists’ work, Constantino concluded that language was a naturally occurring process stemming from the interaction of brain functions and the learning environment. There appeared to be physiologically a critical period for language acquisition spanning from age 2 to puberty. Further, language was developmental and progressed as the brain matures. In the initial acquisition of language, the learner was primarily concerned with meaning and not grammatical precision. Language, she stated, required “rich linguistic input and a positive learning environment” (p. 11).

According to Constantino (1999), research indicated that the acquisition of L2 followed many of the same processes as that of learning a first language (L1). This process too was developmental and required the interaction of physiological and environmental factors. As in L1, the L2 learner initially focused on meaning rather than grammar. Based upon the foundational theories and more recent research, Constantino described four theoretical models for L2 acquisition that provide relevance to L2 reading development and offer implications for the types of program models for ESL students.

As described by Constantino (1999), Cummins’ (1979a, 1979b, 1999) model differentiated between two levels of proficiency in the acquisition of L2. He purported that second language learners first acquire basic interpersonal communication skills (BICS) that include nonverbal and contextual language. For example, BICS allow the learner to interact face-to-face with others during the school day and are less related to

the type of language needed for success in academic learning. This, according to Cummins, required Cognitive Academic Language Proficiency (CALP). Theoretically, as the CALP level increases, there is a greater likelihood that the learner is able to successfully perform tasks with increasing cognitive, academic, and linguistic demands. According to Constantino (1999), the “transitional developmental ‘moment,’ when second language learners acquire CALP was the threshold in Cummins’ model” (p. 16).

Constantino (1999) described Krashen’s (Krashen & Terrell, 1983) model of second language acquisition as one composed of five hypotheses. Comprising the first hypothesis, the Krashen/Terrell model differentiated between acquired language and learned language. Acquired language occurs unconsciously and naturally without conscious attention to grammatical rules. In contrast, learned language is the result of focused attention to the structure and rules of the second language and is typically provided through the teaching process.

Krashen and Terrell’s second hypothesis implied a developmental process of second language acquisition where there was a naturally occurring order of grammatical structures from less to more complex. Not all individuals would develop components of grammar in the same order, but certain structures would occur early in development while others would come later. Constantino (1999) stated that this process has proven true “regardless of the first language of the children learning a second language...” (p. 15). As the learner’s comprehension developed, there was a silent period according to the third hypothesis. During this period, it appeared that children become primarily focused on the sounds of the language in an effort to develop comprehension. This silent period

also occurred during the L1 process as well as L2 and was characterized by limited linguistic expression.

The fourth hypothesis was termed by Krashen and Terrell as the *comprehensible input hypothesis* and implies language is acquired “slightly beyond the current level of acquired competence” (Constantino, 1999, p. 15). Constantino elaborated that when second language learners comprehended communication from others, the content contained structures the learner was ready to acquire. This hypothesis also indicated that language reception must occur before speaking, reading, and writing.

The fifth and final hypothesis was defined as the *affective filter hypothesis* and recognizes the importance of affective factors such as motivation, interest, mood, and self-image in the L2 process (Constantino, 1999). As can be seen, the interaction of the elements from these five hypotheses attest to the complexity of second language learning and according to the Krashen/Terrell model, are essential for successful acquisition.

A third theoretical model described by Constantino (1999) was the Prism Model as proposed by Thomas and Collier (1997). This model is multidimensional and includes academic, cognitive, sociocultural, and linguistic processes that are interactive and interdependent. An integral part of this model is the theory that language and cognitive development are interdependent and as such, when cognitive development in L1 is arrested, there can be negative consequences on cognitive development in L2. This process can be seen when preschool children’s L1 is interrupted by the introduction of L2 through immersion. According to Constantino, Thomas and Collier (1997) believed that all four dimensions of the Prism Model must be addressed by educators if second language literacy was to be successful. One assertion based upon the model is that

academic instruction should be provided in the student's first language for as long as possible while offering a balance of instruction in L2.

A final theoretical model briefly described by Constantino (1999) was the *Social, Linguistic, and Cognitive Processes Model*. A primary contributor to this model was Fillmore (1985) who recognized the importance of linguistic and cognitive processes but emphasized the learning of language in a social context. Constantino noted that Fillmore's perspective is sociological rather than linguistic.

### Educational Program Models

The theoretical models discussed above attest to multidimensional and interactive nature of the L2 process. Taken as a whole, the four models imply that cultural, cognitive, affective, linguistic, sociological, and academic factors must be considered in planning meaningful educational experiences for ESL students and specifically, reading instruction. The actual literature reveals a number of educational models ranging from English only immersion to two-way bilingual programming where both ESL and native speakers are instructed together.

According to Galloway (2003), the most accepted model for literacy instruction in bilingual programs is teaching students in their native language. The conceptual and theoretical basis of this practice implies that there is an interdependence of skills with literacy in L1 transferring and facilitating literacy in the student's second language (Cummins, 1979a). Not all studies have shown significant differences in favor of bilingual programming, however. Slavin and Cheung (2003) reported mixed findings in

their review. They did emphasize that there were no instances where results from English only programs exceeded those from the bilingual programs.

Four definitive works (Constantino, 1999; Collier, 1992; Ramirez, 1992; Thomas & Collier, 2002) provided support for the long-term effectiveness of instructing ESL students in their native language. Ramirez directed a federally mandated longitudinal study following approximately 2,000 elementary Spanish-speaking students for a 4-year period of data collection. The students were well matched on factors such as socioeconomic status, pretests, and preschool experience. The research question focused on determining which of three instructional programs would best serve ESL students. The English-immersion program type used only English to teach the content areas. Early-exit programs were characterized as providing 30 to 60 minutes of L1 instruction per day where primarily, reading skills were taught. Students in the early-exit programs usually began in kindergarten and any instruction in L1 was phased out by the end of 2 years. The late-exit programs were described as providing a minimum of 40% of instruction in Spanish. Students remained in the late-exit programs through their sixth grade year even if they achieved adequate levels of English fluency prior to that time. General findings from this study revealed that learning a second language could require six or more years. Students incurred no adverse effects from being exposed to instruction in L1 and, in fact, the more instruction they received in L1 the greater the likelihood they would eventually achieve at the same levels in L2 as monolingual students.

In examining reading achievement, Ramirez (1992) compared students on standardized measures of reading based upon participation in English immersion and early-exit only programs. He then assessed the achievement of students participating in

three implementations of the late-exit model. As a final analysis, he investigated the achievement growth rates of students in the three program types in comparison to national norms. Ramirez reported no significant differences in reading, mathematics, and language skills between the English immersion and early-exit groups after a period of 4 years when tested in English. Among three late-exit sites, Ramirez found that the site with the most consistent amount of L1 instruction had higher reading scores at the end of sixth grade than the other two sites with apparently less L1 instruction. Ramirez noted variation in the effectiveness and growth rates of reading, mathematics, and English language skills among late-exit programs. He attributed this variation to the proportion of English used for instruction.

When the rate at which students in the English immersion, early-exit, and late-exit programs increased their reading skills were compared to national norms, Ramirez (1992) found all three programs revealed rates as fast as or faster than the norming population. For this aspect of the analysis, growth was measured from kindergarten to third grade. For those students in the late-exit model with the greatest amount of L1 instruction, Ramirez reported their reading growth rate to exceed that of the norming population over a period from kindergarten to sixth grade. Ramirez speculated that if students maintained this growth rate, they would be expected to catch up with their monolingual peers. Of importance, Ramirez observed that students who were “abruptly transitioned” (p. 39) into an English only program appeared to actually lose ground when compared to the norming population.

In a synthesis of research, Collier (1992) critically examined 17 longitudinal studies published since 1980 that focused on four program types. These included the

same three program types described by Ramirez (1992) plus the addition of two-way programs. In two-way programs, ESL and monolingual students were grouped together and instruction was devoted to literacy and fluency in both languages. All studies reviewed by Collier (1992) followed the achievement of students for 4 or more years. Among the four studies examining the achievement for students in two-way programs, all achieved above the 50<sup>th</sup> normal curve equivalent (NCE) on standardized measures of English reading. Similarly, data from five late-exit model studies revealed NCE scores ranging from 44 to 52 on measures of reading. The amount of L1 support, when there was a balance of L2 support, was directly proportional to the students' academic achievement. A key point made by Collier (1992) was the need for a balance of support in L2 rather than instruction solely in L1. The need for L2 support in bilingual programs has also been emphasized by Cummins (1999) who stated that Spanish-speaking "students require a strong English literacy development program in the early grades just as they require a strong Spanish language literacy program" (p. 5).

Constantino (1999) provided a synthesis of research on reading instruction for ESL students in the U.S. to serve as a resource for legislators and educators. As described above, she examined the theoretical foundations of second language development and denoted a basic premise that successful second language development precedes successful reading achievement in L2. Additionally, Constantino examined the difficulties ESL students have encountered in learning a second language, the skills essential for reading instruction, and the program, school, and teaching characteristics prerequisite for ESL students to be successful academically. In determining effective educational programs, Constantino summarized two major findings from her synthesis.

First, research findings indicated a positive correlation between the amount of instruction in the ESL students' first language and their long-term academic achievement in L2.

Further, there appeared to be a positive correlation between the degree of formal schooling in L1 with the rate that ESL students acquire English as their second language.

Second, programs that provided long-term L1 instructional support such as those referred to as late exit and two-way bilingual were shown to promote achievement levels for ESL students that eventually equaled those of monolingual English speaking students.

Programs providing little or no L1 instructional support did not achieve equivalent results. Constantino stated that previous studies denoting little difference among various program models and outcomes measures of English academic achievement were invalid due to their lack of longitudinal focus and little attention to achievement in the early grades.

Thomas and Collier (2002) provided more recent evidence of the positive benefits of instruction in L1. They served as principal investigators in a national report on the effectiveness of schools in addressing language minority students' academic achievement. In a 5-year study extending from 1996 to 2001, the researchers examined records of 210,054 students in school districts located in the northwest, northeast, south central and southeast U.S. The data were described as long-term and represented five school districts. While there were 80 different primary languages represented in the sample, the majority was Spanish-speaking. Findings were reported both from the standpoint of achievement in English on nationally standardized tests, as well as in Spanish. In examining the benefits of total English immersion where parents had refused bilingual educational programming, Thomas and Collier (2002) reported this group

showed significant decreases in reading achievement by fifth grade. Additionally, this group contained the largest number of dropouts. Similar to Collier's (1992) conclusion, as the amount of time of instruction in L1 increased, so did reading scores as students progressed through the grades. Program types described as 90-10 one-way and 50-50 two-way were determined to be the most effective of all programs examined. The 90-10 one-way program provided 90% of instruction in the students' native language through grade 2 with 10% in English. In considering the students' socioeconomic status (SES), the authors concluded that a dual-language program could serve to "reverse the negative effects of SES more than a well-implemented ESL content program by raising reading achievement to a greater degree" (p. 5).

The outcomes from longitudinal studies offered rather compelling evidence of the benefits of academic instruction in L1 on the long-term achievement of ESL students. There are some precautions, however, that should be considered when interpreting these findings. As noted by Collier (1992) and Constantino (1999), methodological issues arose in these studies due to problems controlling treatment variables such as consistent definitions of "bilingual program" and the instructional variation among programs with the same label. Further, longitudinal studies present a challenge to researchers in following the same group of students over time. One major difficulty in the longitudinal approach was maintaining enough of the same students over the course of the study to yield data that could be generalized. Data collected on the basis of cross-sectional analyses could provide a compromise to the longitudinal approach. As explained by Collier (1992), the cross-sectional approach utilizes analyses of all language minority students at each grade level in an effort to acquire enough cases for results that can be

generalized. The author posited that a cross-sectional analysis would likely yield a “more conservative estimate of how a group of students [might] fare in a given program” (p. 189).

A final limitation discussed by Collier was the difficulty in validly interpreting standardized achievement tests for ESL students. For those students evaluated in the earlier grades with limited English skills, their abilities were under estimated when compared to a normative group of native English speakers. Despite these limitations, longitudinal studies did offer positive insights to schools when considering reading achievement and the direction to pursue with bilingual programming.

### Systematic and Explicit Reading Instruction

The studies reviewed above provided support for balanced instruction in L1 and L2 and suggested that such instruction should start at the very beginning of ESL students’ school career. However, for the students who have had 2 to 3 years of traditional English immersion programming with ESOL support, the benefit of then introducing instruction in L1 appears to lack research evidence. For these students, the amount of ESL support may have decreased to a minimal level and, as reported by Collier (1992), their reading achievement may begin to decline in comparison to their monolingual cohorts. The question becomes: what reading strategies are effective in furthering student progress and preventing such a decline? To begin addressing this question, it is first relevant to provide an overview of the current research on explicit and systematic reading instruction.

## The National Reading Panel's Findings

In 1997, Congress requested that the National Institute of Child Health and Human Development (NICHD) create a panel to investigate the research evidence of various methods of reading instruction. The National Reading Panel (NRP) was formed and in 1999 submitted two reports to Congress documenting their findings and recommendations (National Institute of Child Health and Human Development, 2000a). As a basis for launching its investigation, the NRP used the National Reading Council's report, *Preventing Reading Difficulties in Young Children* (Snow, Burns, & Griffin, 1998) to establish the parameters for selecting specific topic areas for investigation. From this report and input from public hearings, the NRP selected the areas of alphabets, to include phonological awareness and phonics instruction, fluency, and comprehension to focus their investigation. Within the realm of comprehension, attention was given to vocabulary and text comprehension instruction. The NRP established specific criteria for studies to include in their investigation and included those that examined reading from preschool age to twelfth grade. When possible, meta-analyses were conducted and if the number of studies were limited, "the NRP made a decision to conduct a more subjective analysis to provide the best possible information about an instructional reading approach or program" (National Institute of Child Health and Human Development, 2000a, p. 5). The NRP specifically did not consider the issue of reading instruction for ESL students. This review will discuss the highlights of the NRP's findings and then turn to an examination of studies addressing systematic and explicit reading instruction for ESL students.

## Phonemic Awareness

The NRP's subgroup report (National Institute of Child Health and Human Development, 2000b) defined phonemic awareness "as the ability to focus on and manipulate phonemes in spoken words" (p. 2-1). Phonemes represent the individual utterances that are combined to form English syllables and words. The NRP identified six tasks that were found in studies to assess and also instruct students in phonemic awareness. These included phoneme segmentation, deletion, isolation, blending, categorization, and identifying the common phoneme sounds in words. On the basis of 96 cases where treatment groups involving phonemic awareness training were compared to control groups with alternate or no instruction, the NRP derived an overall phonemic awareness effect size of 0.86. The NRP noted that effect sizes of 0.80 or above were considered large. Among several conclusions, the NRP summarized that phonemic awareness could be taught and that having the ability to manipulate sounds and letters in words facilitated the reading and writing processes.

In a recent work, Torgesen (2002) stated that one of the most important discoveries in the past 20 years is that difficulties with phonemic awareness placed children at high risk for reading failure. Further, without adequate phonemic awareness skills, phonics instruction is not meaningful. Torgesen (2002) reported that weaknesses in phonemic awareness can be seen across a range of verbal ability. Children with lower intelligence may have early reading difficulties associated with the same core phonological factors as struggling readers with average ability. Gerber et al. (2004) have further suggested the concept of phonemic awareness as representing core cognitive processes and reported that on the basis of increasing empirical evidence, phonological

processing ability may be independent of the language an individual speaks. Researchers (e.g., Torgesen et al. 1999) investigating the merits of phonemic awareness intervention have shown a direct relationship to word identification and subsequently, reading comprehension.

### Phonics

Phonics involves the relationship of letters to sounds and according to the NRP (National Institute of Child Health and Human Development, 2000b), phonics instruction refers to a method of “explicitly teaching students a prespecified set of letter-sound relations and having students read text that provides practice using these relations to decode words” (p. 2-92). The NRP subgroup investigators examining the literature on phonics instruction identified 38 studies containing 66 treatment-control group comparisons. Control group instructional methods involved basal reading programs, whole language, whole word reading, regular curriculum and miscellaneous programs. The results yielded a mean overall effect size of  $d = 0.44$  for phonics instruction. This represented a moderate effect size and according to the NRP, provided solid support that systematic phonics instruction contributed more to reading achievement than other non-systematic methods or no phonics instruction at all. The findings did not indicate that one type of systematic instruction (e.g., synthetic vs. larger unit programs) had superiority over the other. Generally, the key issue was whether the approach was explicit and systematic. In addressing the question of when to begin phonics instruction, the NRP’s analysis indicated that phonics instruction was more effective when introduced in

kindergarten or first grade, and before children had learned to read independently.

Similarly, phonics instruction had a significant impact on the reading comprehension and spelling skills of younger students but the effect decreased for older students.

As noted above, the NRP specifically did not address the issue of reading instruction and second language learners. However, the phonics subgroup investigators did examine studies supporting phonics instruction for students in kindergarten and first grade who were at risk and those who were identified as disabled readers. For low achieving readers in grades 2 through 6, systematic phonics instruction appeared to not have significant impact on their reading achievement. The NRP suggested that for this group, perhaps the phonics instruction was not intense enough, there were too few cases to yield reliable results, and/or their difficulties were due to other problems such as comprehension deficits.

The NRP's report on phonics instruction was not without its critics. Camilli, Vargas, and Yurecko (2003) argued that the NRP's methodology was inadequate in examining phonics instruction. Camilli et al. reexamined the same 38 studies reviewed by the NRP. In the process, they deleted one study and added three additional ones to yield a smaller but still significant effect size ( $d = 0.24$ ) for phonics instruction. However, they found that the effect size tripled when phonics instruction also included systematic language activities, and individual tutoring. The NRP's meta-analysis had not shown individual phonics tutoring to be favored over small group or larger group instruction. The Camilli et al. findings offered another possible explanation for the NRP's determination that low achievers did not benefit substantially from systematic phonics

instruction. Further, Camilli's et al. results may also have implications when considering reading instruction for the older student learning a second language.

### Fluency

The NRP subgroup report described fluency broadly as the ability to read with speed, accuracy, and expression. In justifying the inclusion of fluency in their analysis, the NRP reported evidence from a U.S. Department of Education report (Pinnell et al. 1995) indicating dysfluency among a national sample of fourth graders, and secondly, showing fluency was closely related to reading comprehension. In their work, Pinnell et al. described fluency as the automaticity of reading where the reader does not need to labor with decoding but can instead, focus on meaning. To examine the research evidence on fluency, the NRP divided their investigation into two broad avenues relevant to the methods of increasing fluency. One examined guided reading where essentially, fluency is attained through repetition. The other considered research where students were given less structure and were generally encouraged to read more.

The NRP's findings indicated moderate support for guided reading approaches as a means of increasing fluency. Oral reading practice appeared to have the greatest influence on measures of reading speed, accuracy of oral reading, and word knowledge. In some studies, oral reading practice was found to have an impact on reading comprehension and overall reading ability. According to the NRP, in several studies, the impact was "actually quite high" (p. 3-18). The NRP suggested that the influence of oral practice could vary dependent on the reading level of the individual with higher readers

reaping greater comprehension benefits than those at lower levels where word recognition might be enhanced instead.

The NRP's second synthesis on fluency where students were encouraged to read more often failed to show support for this method and placed schools' practice of encouraging voluntary reading in a questionable light. The conclusions were that guided oral reading where readers read passages multiple times with feedback could improve fluency and, in turn, a variety of other reading skills. Further, this type of fluency procedure appeared to help students through fifth grade and older students with learning difficulties.

To provide a balanced view of instruction where students are encouraged to engage in more voluntary reading, it is relevant to note that Krashen (2001), for example, has voiced strong disagreement with the NRP's findings on free reading involving sustained silent reading. He stated that the NRP failed to include relevant studies and misinterpreted some that were included. By expanding the number of studies, Krashen found that students who engaged in sustained silent reading performed as well as or better on measures of reading comprehension than control students in 50 out of 53 comparisons. Krashen noted that the NRP did not include studies lasting longer than one year and that the benefits of recreational reading appeared to increase over time. Further, by reinterpreting only those included by the NRP, Krashen claimed recreational reading still had positive benefits on literacy.

## Comprehension

The NRP subgroup addressed the subject of reading comprehension by analyzing studies on vocabulary instruction, text comprehension, and teacher preparation for the instruction of reading comprehension (National Institute of Child Health and Human Development, 2000b). For the purposes of the present study, the NRP's findings on vocabulary instruction and text comprehension appeared relevant.

## Vocabulary

For vocabulary instruction, the NRP identified a total of 47 studies. However, due to the variations in methodology and conceptual bases of vocabulary instruction, a meta-analysis was not possible. As a result, the NRP provided a review of the "implicit evidence" (p. 4-3) of vocabulary instruction with attention given to the most recent meta-analyses found in their search. Five methods of instruction were identified based upon the studies reviewed. These included explicit instruction where students were given vocabulary words to learn, implicit instruction where incidental learning was involved, multimedia methods, capacity methods where vocabulary was increased by making reading automatic, and methods where students were taught to make associations between words they knew and new ones.

Key findings from the NRP's review of the vocabulary literature included the positive relationship of vocabulary instruction to reading comprehension; vocabulary could be acquired through incidental learning, positive gains in vocabulary were attainable through the use of computer instruction, and there was a need for multiple

exposure to vocabulary instruction over time. The NRP noted that restructuring vocabulary tasks where, for example, easier words were substituted for more difficult ones could increase vocabulary and comprehension skills and seemed most appropriate for at-risk students and low achievers.

### Text Comprehension

The NRP's analysis of 203 studies on text comprehension resulted in 16 methods of instruction with eight of those providing scientific evidence that they actually improved reading comprehension. When the eight methods were considered in combination, gains were shown on standardized measures of comprehension. This approach was in contrast to explicit reading comprehension instruction where students were taught cognitive and reasoning strategies for increasing their comprehension of text. Similarly, the NRP did identify comprehension monitoring as a research supported method where the reader learns to perceive when there is a lack of comprehension and takes steps to increase it. Other specific strategies identified by the NRP included questions posed by the teacher, questions posed by the reader while reading, summarization of content, the use of graphic organizers, cooperative learning, analysis of story structure, and the use of multiple strategies. The NRP emphasized that text comprehension could be addressed best by teachers in "naturalistic settings" (National Institute of Child Health and Human Development, 2000b, p. 4-6) with the flexible use of multiple comprehension strategies as they interact with students over text. Perhaps a key point implied from the NRP's conclusions was the concept of comprehension addressed

flexibly and naturally without the need for explicit and structured instruction on the development of cognitive strategies.

### Explicit Reading Instruction and the ESL Student

As noted above, in their systematic review of the literature, the National Reading Panel (NRP) did not address the subject of reading and second language learners. According to Slavin and Cheung (2003), researchers have given attention to whether these students acquire reading skills in the same manner as their monolingual cohorts or if there are other processes involved. These authors stated that the same factors identified by the NRP have been associated with reading success for ESL students, whether the students have been taught in their native language or not. However, in their review of the literature, Slavin and Cheung found only a “handful of studies” (p. 37) that met criteria for inclusion in their analysis. Minimum requirements included the use of experimental and control groups and treatment for at least 12 weeks. Based upon their search, 11 studies were identified that addressed reading skills for ESL students in kindergarten or first grade. Generally positive results were noted for intervention programs such as Success for All (Slavin & Madden, 1999), Direct Instruction (Becker & Gersten, 1982; Gersten, 1985), and a systematic phonics program (Stuart, 1999). For the upper elementary grades, 10 studies were located that also showed positive effects on reading for programs involving cooperative learning strategies (Calderon, Hertz-Larowitz, & Slavin, 1998), English vocabulary intervention (Carlo et al., 2004), instruction in English oral language skills (Perez, 1981), and a phonetic program provided to very low

achieving ESL students (Denton, 2000). At the secondary level, one study (Shames, 1998) was found to show positive gains for ESL students in reading comprehension through cooperative learning and direct instruction approaches. Slavin and Cheung concluded that effective reading programs for ESL students were likely to be similar to those for monolingual English speakers if there are adaptations to account for their limited English proficiency. They also stated their analysis showed “how much remains to be done on effective reading programs for English language learners” (p. 37). Relevant to the present study is current research that addresses specific areas identified by the NRP in relation to ESL students.

#### Phonemic Awareness and Phonics

Gerber et al. (2004) stated that the research on systematic and explicit reading instruction for ESL students, whether they were or were not experiencing difficulties, appeared limited. However, there has been increasing attention given to the benefits of instruction in phonemic awareness for second language students.

Two separate studies were identified where treatment involved systematic and explicit English reading instruction, including phonemic awareness, but treatment groups differed in the initial language in which students were taught reading. In a work by Gunn, Biglan, Smolkowski, and Ary (2000), Hispanic students who participated in the treatment group apparently were never taught to read in Spanish. According to the authors, the study was conducted because little was known about the usefulness of a systematic approach to teaching reading to Hispanic students and “direct tests of the efficacy of

teaching reading in English [had] not been conducted” (p. 91). The authors reported strong evidence that many students with reading difficulties had deficits in phonological awareness, rapid naming, and phonological recoding. For their independent variable, Gunn et al. (2000) provided supplemental English reading instruction to groups of Hispanic and non-Hispanic students beginning in kindergarten and continued the instruction through third grade. Specifically, the students were tutored in phonemic awareness, sound-letter correspondence, blending words, and, in turn, were given practice reading decodable text. After testing over three different time periods, their findings revealed that students in the treatment group differed significantly in word attack, word identification, oral reading fluency, comprehension, and vocabulary when compared to controls. There were no differences based upon students’ level of English fluency, gender, or grade. More specifically, Hispanic students who, at the beginning of the study spoke little or no English, benefited from systematic phonics instruction as much as students with greater English proficiency. In a follow-up study, Gunn, Smolkowski, Biglan, and Black (2002) found that the students who had received supplemental instruction continued to show their acquired gains in reading fluency and word attack one year later.

In another work involving systematic instruction in phonics, Denton, Anthony, Parker, and Hasbrouck (2004) tutored Spanish-speaking bilingual students in grades 2-5 three times per week for approximately 3 months. In contrast to Gunn et al. (2000), both the experimental and control groups had received reading instruction in English and Spanish. In addition to phonics, the students also received vocabulary and comprehension instruction. Using repeated-measures mixed analysis of variance for data analysis, the

researchers found significant gains in word identification as related to phonics instruction when the tutored students were compared to a matched group of non-tutored classmates. Gains were not seen in reading comprehension. On a measure of English fluency, the treatment and comparison groups' mean scores ranged from 2.69 to 3.76 on a scale of 1 to 5, with 5 representing the highest level of proficiency. Slavin and Cheung (2004) in their review, reported that consistent positive results for ESL students were found for the use of systematic phonics instruction. They specifically identified the Success for All and Direct Instruction programs as examples.

Leafstedt (2002) examined the issue of language of instruction and the relationship to reading decoding skills and concluded that neither Spanish nor English as the language of instruction significantly predicted English pseudoword decoding skills. The use of pseudowords requires the reader to rely more heavily on phonological decoding rather than vocabulary knowledge. Leafstedt did find that language of instruction was significantly correlated with Spanish pseudoword decoding. However, when considering the identification of real words, there was a significant relationship with language of instruction. These findings suggested that vocabulary knowledge in English would facilitate English word identification. Similarly, Gottardo (2002) reported that the strongest predictors of word identification in English were L2 vocabulary knowledge, L1 and L2 phonological processing, and consistent with other studies reviewed earlier, L1 reading. Gottardo noted that more recently, the literature addressing the effects of oral language proficiency seemed to show a stronger relationship with reading comprehension than word recognition.

Research attention has also been given to phonemic awareness training in L1 with evidence that skills transfer to English reading. Leafstedt (2002) and other researchers (e.g., Durgunoglu, 2002; Hudson, Smith, & Smith, 2001; Nelson, 2003) have emphasized the importance of cross-language transfer of phonological skills when students move from reading in L1 to literacy in L2. Duran, Shefelbine, Carnine, Maldonado-Colon, and Gunn (2003) reported on one study (Duran & Carnine, 1999) of phonological awareness training in Spanish where students subsequently were found to be better English decoders than students who had not received such training in Spanish. Gerber et al. (2004) conducted a longitudinal study where at risk kindergarten ESL students were identified on the basis of teacher input and bilingual phonological skills. By using a direct instruction method to teach phonemic awareness skills in Spanish, students essentially caught up with peers who did not have initial reading difficulty on measures of English word reading. The authors noted that their treatment approach was consistent with the multitiered model promoted by Reading First.

### Fluency

While there has been increasing attention given to phonemic awareness and phonics, there appeared relatively little in the literature addressing reading fluency for ESL students. In her review of practices in the development of English literacy for ESL students, August (2003) stated that these students may have less opportunities to practice reading aloud at home than non-ESL students since their parents may not be English proficient. Further, fluency appears enhanced if students understand the words they are

reading, a further limitation of limited English proficiency. August did report one study where researchers examined the use of assisted reading to improve the fluency of 12-year-old Spanish speaking students (Van Wegenen, Williams, & McLaughlin, 1994). Students read a passage silently while listening to a recording of the teacher's reading, read aloud, read the passage silently three times while listening to the recording, and then read aloud again. The findings showed increased reading rate when measuring words per minute, a reduced number of errors, and improved reading comprehension when compared to baseline measures. August did note that written activities were included that focused on vocabulary and "understanding the significance of each word" (p. 18). When assessing the reading comprehension outcomes, it would seem unclear whether the repetitive reading activities and/or vocabulary training were instrumental. What is obvious, is the need for more investigation into the effects of reading fluency training on ESL students' achievement.

### Vocabulary

As noted by Slavin and Cheung (2003), effective reading approaches for monolingual English only students have much to offer students learning English but must be considered with adaptations. It would seem that a primary consideration would be the limited English vocabulary of ESL students and its relationship to reading achievement. Several researchers (e.g., Brisbois, 1995; Fisher & Cabello, 1981; Grabe, 1991) have found support for the importance of vocabulary instruction in the development of ESL students' reading progress and particularly, written language comprehension. August,

Carlo, Dressler, and Snow (2005) reported that data collected from a cross section of fourth grade Spanish speaking students in California, Virginia, and Massachusetts revealed limited English vocabulary development both in terms of breadth and depth. Given the important relationship between vocabulary and reading comprehension, August et al. noted the poverty of experimental and quasiexperimental studies that had been done on English vocabulary instruction with elementary ESL students. In her review, August (2003) identified a number of practices that appeared effective. She broadly differentiated between incidental vocabulary acquisition and that learned through reading text. For example, Laufer (2001) found that requiring students to use target words such as in sentences, was more effective in increasing vocabulary knowledge than when students only read or heard the words. August et al. (2005) reported on several studies where vocabulary intervention strategies were assessed with many of them adaptations of interventions used for English only students. The authors noted that there were several strategies that appeared particularly important for ESL students. First, it was recommended that educators take advantage of the students' first language and assist them in identifying cognates that the two languages shared. This, in turn, could help increase meaningful vocabulary. Second, teachers should ensure that ESL students understand the meaning of basic words and particularly, words that appear often in texts and instructional materials. Third, the authors stressed the importance of review and practice through the use of both teacher-directed and student-directed activities. This was reminiscent of the National Reading NRP's recommendation for multiple exposure of vocabulary over time.

Essentially, the literature on ESL students and their vocabulary development in L2 appears to imply a direct relationship with reading comprehension and the critical need to address vocabulary skills in conjunction with decoding. However, it appears important to note that intervention may be further complicated when the student has limited proficiency in L1. As noted by Collier (1989), introducing English before the student's development in L1 has been completed can disrupt the progress of L1 and the student may then have limited vocabulary in both languages.

### Text Comprehension

August (2003) attempted to address the issue of text comprehension for ESL students in view of the National Reading NRP's findings and recommendations. One example of the research cited by August focused on the hypothesis that limited vocabulary was related to limited text comprehension (Hakuta, Butler, & Witt, 1999). August's review indicated that vocabulary could be a significant factor in explaining the variation in reading comprehension between weak and strong readers. Other outcomes from this study included the importance for readers to attend to syntax clues, use metacognitive strategies, decrease the number of miscues, which can change the meaning of sentences, and the fact that L2 readers needed more "rich context" (p. 26) to comprehend well.

Another factor identified in August's (2003) review important for comprehension was the structure of language. L2 readers may focus on non-essential items in text due to differences between L1 and L2 and, as a result, may require explicit instruction in the

structure of L2. August specifically noted that word order variations, the complexity of noun phrases and clause formations may serve to mislead the ESL student when reading, particularly at the beginning stages of reading.

A final area that appeared particularly relevant to reading instruction and comprehension for ESL students as identified by Fitzgerald (1995) was the concept of schemata. In her review, she found that the reader's schemata could directly affect comprehension and recall of content. Essentially, the more familiar the reader was with the content of text and the more consistent passages were with the reader's own culture, the better the comprehension. Such findings would have direct implications for teachers of ESL students and the need to ensure that the context of the reading content is familiar and understood.

August (2003) identified a number of practices that appeared hopeful for addressing the ESL students' comprehension with most of them consistent with recommendations made by the National Reading NRP on text comprehension strategies. Specifically, August noted that in addition to vocabulary development, scaffolding instruction, the generation of inferences, using cognate strategies, and using multiple comprehension strategies had received research support. Duran, Shefelbine, Carnine, Maldonado-Colon, and Gunn (2003) stated that reading comprehension for ESL students should be approached in the same manner as teaching decoding, systematically and sequentially. They noted that a comprehension program should have elements involving structured vocabulary training, sequential introduction of specific comprehension skills, and passage reading where the various skills could become integrated.

## Conclusions

The research reviewed here offered rather substantive evidence of programmatic differences in reading achievement for ESL students. Significantly positive benefits were seen for balanced instruction in both L1 and L2 when compared to English only immersion programs. Students who have participated in literacy programs where there was a combination of instruction in their native language and English appeared to have the best opportunities for catching up with monolingual students in reading achievement and in maintaining their level of skills throughout their school careers. When teaching reading to ESL students, there was evidence that many of the same practices appropriate for monolingual learners were appropriate if, according to Slavin and Cheung (2004), adaptations and modifications were provided. Findings by researchers such as Gunn et al. (2000, 2002), offered some support that for those ESL students who have not had benefit of reading instruction in L1, there still could be development of effective decoding skills in L2 through explicit instruction. Further, with a balance of vocabulary instruction, the use of multiple comprehension strategies, and emphasis on oral language, reading comprehension could be enhanced. Such findings suggested that any reading program for ESL students should include emphasis on the phonological processes, vocabulary knowledge, and oral language competency. Research where treatment variables included all of the combined skills emphasized by Reading First (i.e., phonics, fluency, phonological awareness, vocabulary, and comprehension) appeared absent and suggested a need for investigation of the usefulness of such a combined approach for ESL students who have participated only in an immersion program.

### Collective Teacher Efficacy

Teacher self-efficacy has received considerable research attention with theoretical underpinnings based upon Bandura's (1977, 1986, & 1997) social cognitive theory and Rotter's social learning theory (1966). In social cognitive theory, self-efficacy refers to "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p. 2). Teacher efficacy refers to the belief that one can engender student learning and is a group of self-efficacy beliefs (Ross, Hogaboam-Gray, & Gray, 2003) Teachers' efficacy beliefs have been related to student achievement, student attitudes, teacher attitudes, teacher stress, and others connected to both teacher and student factors (Brouwers & Tomic, 2001). Brouwers and Tomic stated that teacher efficacy is "a critical variable in studying many educational concerns" (p. 78). Ross et al. reported that teacher efficacy contributes to student achievement because teachers with high efficacy beliefs appear more motivated to facilitate student learning, provide more attention to students with lower ability, foster student autonomy, and can change how students perceive their own abilities.

Bandura (1997) has specifically addressed collective efficacy as a means of assessing the efficacy beliefs among groups. Accordingly, he has defined collective efficacy as "a group's shared belief in its conjoint capabilities to organize and execute the courses of action required to produce given levels of attainments" (p. 477). Bandura (1997) noted that there is a two-way interaction between individual efficacy beliefs and those of the group. What emerges is a "property that is more than the sum of the individual attributes" (p.477). Ross, Hogaboam-Gray, and Gray (2003) identified sources of teachers' collective efficacy beliefs and included prior achievement, which acts to

generate a sense of mastery for teachers. The authors reported that mastery experiences tended to be the most powerful in contributing to and predicting collective teacher efficacy. Other sources noted by Ross et al. were teacher collaboration, affective states, and social persuasion, which consist “of organizational members persuading other members that they constitute an effective team” (p. 9). At the time of their study, Ross et al. stated that no study had investigated the role of the school principal in contributing to collective teacher efficacy but did report that supportive and transformational leadership was related to teacher efficacy.

Goddard, Hoy, and Hoy (2000) observed that collective teacher efficacy had received relatively little research attention. In their work, a conceptual model of collective efficacy was developed for use in schools. They established that both group competence and task analysis were integral components of collective efficacy and both were significantly related. They concluded that their model represented an extension of individual teacher efficacy. In turn, the authors created a measure of collective teacher efficacy and subsequently, found it to have high reliability and validity. Using the instrument in an urban school district, they found that differences in collective teacher efficacy were positively associated with variations in student achievement in reading and mathematics. Thus, teachers’ beliefs in their collective ability to effect academic progress could significantly influence a school’s overall achievement levels. Further, such beliefs could mediate the effects of student SES and students’ prior achievement levels. The authors posited that collective efficacy had a greater effect on student achievement than students’ SES. The authors operationalized student SES as either (1) participants in the free or reduced cost lunch program or (0), all others. In another work, Goddard, Logerfo,

and Hoy (2004), assessed high school teachers' collective efficacy and found it to be a significant predictor of twelfth grade student achievement in all of the curricular areas evaluated by the state.

Ross, Hogaboam-Gray, and Gray (2003) questioned the extent to which school processes and student achievement contributed to collective teacher efficacy. School processes were identified as variables representing school cohesion and support, and the process by which teachers shared in school decision making. Their sample included 141 schools in Canada with 2170 teachers responding. The authors used a 14-item variant of the original scale developed by Goddard et al. (2000). Their findings were consistent with others (e.g., Goddard, 2002) that prior student achievement was predictive of collective teacher efficacy. However, they found school processes to be greater than prior achievement. Specifically, school-wide collaboration, empowering school leadership, shared school goals, and *fit of plans with school needs* had the strongest influences on collective teacher efficacy. Ross et al. explained that one factor affecting the lower influence of student achievement might have been the teachers' belief that the mandated achievement testing results lacked validity.

In a later but somewhat similar work, Dale (2005) examined the relationship between collective teacher efficacy and the variables of teacher trust of teacher, teacher trust of principal, teacher-teacher collaboration, teacher principal collaboration, SES, prior academic skill, and school level. Significant predictors of collective teacher efficacy included teacher-teacher trust, prior academic skill, and teacher-teacher collaboration. Interestingly, the variables of teacher trust of principal and teacher-principal collaboration were not found to be significant predictors.

### Measurement of Collective Teacher Efficacy

According to Bandura (1997), collective efficacy beliefs could be assessed through amassing individual group members' beliefs about their own personal capabilities for the skills they employ as a group. In a second approach, collective efficacy could be considered on the basis of members' appraisals regarding the whole group's capability.

Tschannen-Moran, Hoy, and Hoy (1998) noted in their review of teacher efficacy that while collective efficacy's importance appeared to warrant research attention, there were no consistent measures to assess it. Since then, there have been attempts to quantitatively analyze schools' collective teacher efficacy. The development of such a scale has received particular attention from Goddard et al. (2000) and Goddard (2002). Goddard et al. determined that their Collective Teacher Efficacy Scale (CTES) produced reliable and valid scores. They initially created a 21-item scale and conducted pilot tests. Findings revealed a significant correlation ( $r = .54$ ,  $p < .01$ ) between personal teacher efficacy and collective teacher efficacy. Additionally, they found the CTES to have high internal reliability ( $\alpha = .96$ ).

Goddard (2002) later sought to prove evidence of the validity of scores on a shorter form of the 21-item scale. He concluded that a 12-item scale was more 'theoretically pure' (p. 97) than the earlier scale. Similar to the longer scale, items fell within two broad factors involving group competence and task analysis, and were either negatively or positively worded. A significant correlation ( $r = .983$ ) was found between the short and long forms of the scale. In correlating the scale with mathematics

achievement, Goddard concluded that the short form of the scale was a positive predictor of variability in achievement among schools assessed.

Collective teacher efficacy has been shown to have a positive relationship with student achievement. It would seem important to consider this factor when attempting to understand a phenomenon as complex as reading achievement. The relationship of individual teacher efficacy to teaching practices and strategies appeared well addressed in the literature (Dukes, 2000; Guskey, 1987; Starko & Schack, 1989; Wheatley, 1998). However, what appears lacking in the research is the relationship between schools' collective teacher efficacy and innovations in teaching methodology and specifically, when it has been adopted by an entire school, as is Reading First.

## CHAPTER THREE

### METHODOLOGY

#### Research Design

The design of this study was causal-comparative and employed a nonequivalent (pretest and posttest) control group design (Creswell, 2003). This design appeared most appropriate for addressing questions about the statistical significance of differences in reading achievement between two naturally occurring groups that cannot be randomly assigned to experimental and control groups. Subjects for the two groups were matched in an effort to limit as much as possible the influence of extraneous variables including level of English proficiency, the impact of prior grade retention, literacy in the students' heritage language, socioeconomic status, participation in an ESOL program, and participation in special educational programming.

#### Participating Schools

Two schools were identified in a Southwest Florida school district that were as similar as possible in socioeconomic status (SES), percentage of students whose home language was English, percentage of Hispanic students, and equivalency of reading curricula used in addition to Reading First. Additionally, after school tutoring opportunities and reading instructional approaches were considered in matching the two schools. SES determination was derived from the percentage of students who received free or reduced cost lunch. The major difference was that one school was designated as a

Reading First/Title I school while the other was not. Within the school district, all elementary schools receiving Title I funds had been designated as Reading First. These two schools were termed Elementary A as the Reading First school, and Elementary B as the non-Reading First school. Specific demographic data from the two schools are provided in Table 1.

Reading First addresses reading instruction for students in grades K-3 by primarily funding staff development, resource materials, and ongoing assessment (Just Read, Florida!, n/d). Reading First does not attempt to remediate students in small groups or provide any instruction outside the regular classroom setting. Teachers undergo a week long training session to learn the research basis and instructional methods characteristic of the Reading First program. Reading First does not recommend specific texts but trains teachers to use those adopted by their school districts in instructing the five major content areas. As a part of the training, teachers are provided strategies to use with limited English proficient students (Just Read, Florida!, 2002).

At Elementary A, the five major areas of reading instruction emphasized by Reading First were provided to all students in grades K-5. Further, students in grades 4 and 5 were assessed four times during the school year with the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2001) as were the students in grades K-3. Elementary B employed reading materials and texts adopted by the school district and did not assess students using the DIBELS. Both schools participated in the district's standardized achievement testing program where a group reading achievement test was administered in May of 2004 and 2005, and the Florida Comprehensive

Assessment Test (FCAT) was given in March 2005. Students were also tested in the areas of writing and science but results were not relevant to the questions of this study.

Table 1

Demographics of Elementary A and Elementary B Schools

	Elementary A (n = 1053)		Elementary B (n = 933)	
	Reading First		Non-Reading First	
	n	%	n	%
ESL	228	21.65	154	16.52
Hispanic	591	56.13	322	34.54
Free/Reduced Lunch	798	75.78	449	48.14
Exceptional Students	184	17.47	89	9.57
<u>Home Language English</u>	426	40.55	568	60.83

Note: Data were collected from school district's web site during August 2004.

### Participants

The initial step in subject selection for participation in this study was a request that principals at each school release student identification numbers for the total limited English population. By only using identification numbers, the students remained anonymous and their identities were protected. Strict confidentiality of student identification numbers was maintained and only identification numbers were associated with any test data. Once obtained, the student identification numbers were then used to access district electronic databases to begin subject selection.

A sample of 34 students was available from Elementary A with an identical number located at Elementary B who met criteria for selection. Subjects were selected who met the following criteria: (a) were limited English proficient, (b) participated in ESOL during the 2004-2005 school year, (c) had no prior grade retentions, (d) were not identified as learning disabled, language impaired, or emotionally handicapped (e) had been tested with the Group Reading Assessment and Diagnostic Evaluation (GRADE) (Williams, 2002) during May 2004 and May 2005, (e) were students in grades 2 through 5 in either Elementary A or B, (f) first language was Spanish, and (g) participated in the district's free or reduced cost lunch program. Students meeting all criteria but who had not attended the Reading First school for the entire year were excluded. ESOL personnel reported that two of the third grade Reading First students and three of the fifth grade non-Reading First students had attained some degree of Spanish literacy skills. The extent, however, of their Spanish literacy skills was not assessed as a part of this study. Kindergarten and first grade students were not selected due to the scarcity of 2004 GRADE test scores for this group. Limited English proficiency was defined in accordance with the Bilingual Verbal Ability Tests (BVAT) (Muñoz-Sandoval, Cummins, Alvarado, & Ruef, 1998) for limited proficiency specifying a Cognitive Academic Language Proficiency (CALP) level of 2.5 to 3.5. Students in both groups had achieved this level of English fluency by the beginning of the 2004-2005 school year. Students in both groups were participants in regular mainstream classrooms and received services through the ESOL program during the 2004-2005 school year. In both schools, ESOL was an English only program and provided inclusion tutoring by paraprofessionals

under the direction of the classroom teacher. Table 2 provides a summary of student demographics for each of the two groups.

Table 2

Student Demographics

	Elementary A (n = 34)		Elementary B (n = 34)	
	n	%	n	%
Gender				
Female	16	47	18	53
Male	18	53	16	47
Grade				
Second	8	24	8	24
Third	8	24	10	29
Fourth	13	38	7	21
Fifth	5	14	9	26
Total	34	100	34	100

Participants for the second part of the study included teachers at Elementary A and B. Teachers at each school were asked to complete the Collective Teacher Efficacy Scale – Short Form (CTE) (Goddard, 2002) (Appendix A) and a brief teacher questionnaire (Appendix B). At Elementary A, a total of 38 of the 53 teachers (71.7 %) completed the CTE and 37 completed the questionnaire. At Elementary B, 30 of the 50 teachers (60 %) completed the CTE. Among this group, 28 returned the questionnaire. Table 3 provides demographic information for teachers at each of the two schools based upon the questionnaire. As can be seen, the majority of teachers in Elementary A had

received Reading First training while the majority of Elementary B had not. A total of 8 teachers at Elementary A had not completed Reading First training. Among the 8, 3 were at fifth grade, 2 at fourth grade, and one each from grades K, 1, and 3. Students in these classes were still expected to achieve skills in all reading areas emphasized by the Reading First program and were tested along with all other students on the DIBELS four times during the school year.

Table 3

## Teacher Demographics

	Elementary A (n = 37)		Elementary B (n = 28)	
	n	%	n	%
<b>Gender</b>				
Female	35	95	28	100
Male	2	5	0	
<b>Years Teaching</b>				
First Year	9	24.3	2	7.1
2-5	12	32.4	8	28.6
6-10	10	27.0	9	32.1
11-20	4	10.8	3	10.7
20+	2	5.4	6	21.4
<b>Highest Degree Earned</b>				
BA/BS	26	70.3	14	50.0
MA/MS	11	29.7	14	50.0
<b>Reading First Trained</b>				
Trained	29	78.4	2	7.1
Not Trained	8	21.6	26	92.9
<b>Years at Current School</b>				
First Year	12	32.4	5	17.9
2-5	20	54.1	23	82.1
6-10	2	5.4	0	
10+	3	8.1	0	

Note: Elementary B opened at the beginning of the 2002-03 school year.

## Procedures

### Instrumentation

To determine the equivalency of reading levels between the Reading First and non-Reading First groups, results were available from informal reading inventories on 25 of Elementary A's 34 students and 27 of the 34 students from Elementary B. These reading measures were administered during the 2002-2003 school year. To establish initial reading levels for students, the district utilized the Developmental Reading Assessment (DRA) (Beaver, 1997) and to a lesser degree, the Flynt-Cooter Reading Inventory (Flynt & Cooter, 1998). Scores for students in this study were available in the district's electronic database. Reading levels from the Flynt-Cooter could be converted to DRA equivalences using a conversion chart provided by the school district. For purposes of this study, all scores were converted to DRA levels to make data interpretation less cumbersome. DRA levels for each group are summarized in Table 4. Levels range from A to 44 and are not equally spaced. The DRA was designed for use in grades K-3 and is administered individually by the teacher. As the student reads selected text from the DRA, the teacher completes a running record and the DRA Observation Guide. Results from the DRA allow the teacher to monitor the students' progress and can be used for grouping students. The DRA was initially developed and field tested by primary teachers in the Upper Arlington City School District in Ohio between 1988 and 1996. According to the author, "procedures, forms, and assessment texts have changed over the years in response to teachers' feedback and suggestions" (p. 7).

Table 4

## Developmental Reading Assessment Levels from the 2002-03 Academic Year

Elementary A (n = 25)				Elementary B (n = 27)			
Grade 2 n = 5				Grade 2 n = 8			
DRA Levels	n	%		DRA Levels	n	%	
A-14	5	100		A-14	7	87.5	
16-18	0			16-18	1	12.5	
Grade 3 n = 6				Grade 3 n = 10			
DRA Levels	n	%		DRA Levels	n	%	
A-14	5	83.3		A-14	6	85.7	
16-18	1	16.7		16-18	1	14.3	
Grade 4 n = 9				Grade 4 n = 7			
DRA Levels	n	%		DRA Levels	n	%	
A-14	1	11.1		A -14	5	71.4	
16-18	2	22.2		16-18	1	14.3	
20-24	2	22.2		20-24	1	14.3	
28-30	3	33.3		28-30	0		
34-40	1	11.1		34-40	0		
Grade 5 n = 5				Grade 5 n = 5			
DRA Levels	n	%		DRA Levels	n	%	
A-14	0			A-14	1	20.0	
16-18	0			16-18	3	60.0	
20-24	4	80.0		20-24	1	20.0	
28-30	1	20.0		28-30	0		

Note: DRA Levels: A – 14 Readiness to Primer Level, 16 – 18 First Grade , 20 – 24 Second Grade, 28 – 30 Third Grade, 34 – 40 Fourth Grade, 44 Fifth Grade

The Flynt-Cooter was described by the authors as an informal reading inventory for assessing reading levels from preprimer to grade 12. Teachers typically administer the Flynt-Cooter to determine reading levels, placement, and to establish appropriate reading materials needed for instruction. The inventory consists of an interest/attitude

questionnaire, sentences to establish initial passage selection, the reading passages, and assessment protocols. To assess reading skills, miscues are analyzed by the teacher for each level.

The Group Reading Assessment and Diagnostic Evaluation (GRADE) (Williams, 2002), was administered by the school district during May 2004 and May 2005 to all students in the study. The two GRADE administrations served as pre- and posttests. The GRADE is a group measure given by teachers in their classrooms. In addition to a Total Test Standard Score, the GRADE provides separate scores for varying combinations of Word Reading, Word Meaning, Vocabulary, Sentence Comprehension, Passage Comprehension, and Listening Comprehension subtests depending on grade level. The Standard Score has a mean of 100 and a standard deviation of 15. According to the test's technical manual, reliability coefficients obtained by using an alternate form and test-retest procedures were in the .90 range. Concurrent and predictive validity have been established using a variety of other standardized tests.

To provide a measure of English fluency, the Bilingual Verbal Ability Tests (BVAT) (Muñoz-Sandoval et al., 1998) were administered to all participants by school personnel. The BVAT is used to determine initial eligibility for the district's ESOL program and, in turn, is given each year until the ESL student exits the program. Participants in the present study were given the BVAT at the end of the 2004 school year or at the beginning of the following school year. Administration is individual and typically, is given by a bilingual examiner. In addition to Standard Scores, results yield Cognitive Academic Language Proficiency (CALP) levels ranging from 1 for Negligible English to 5 for Advanced proficiency. The CALP measure is based upon the conceptual

framework established by Cummins (1999). The publisher reported a median alternate form and alternate procedure reliability coefficient of .84 for students in grades one through 12. Validity was reportedly established with a variety of criterion measures. Consideration of the BVAT scores in this study appeared supported by other research (Durgunoglu, Nagy & Hancia-Bhatt, 1993; Geva, 2000), which suggested that oral language proficiency is positively related to reading comprehension.

The Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2001) was developed to measure literacy skills involving letter naming, reading fluency, phoneme segmentation fluency, and phoneme onset fluency. The students' grade level determines which subtest or combinations of subtests are given. For Reading First participants, the DIBELS was given individually four times during the school year with one administration in the fall, two in winter, and one in spring of 2005. Results from the spring 2005 administration were analyzed as a part of this study. Students in grade 2 were given the Nonsense Word Fluency and Oral Reading Fluency subtests. Grades 3-5 were given only the Oral Reading Fluency subtest. According to Gunn, Biglan, Smolkowski, and Ary (2000), reliability of the DIBELS Oral Reading Fluency subtest has been established within a range of .92 to .97 using test-retest and alternate forms methods. When correlated with other standardized reading measures, correlation coefficients ranged from .73 to .93. Criteria have been established for cut-off scores that allow interpretation on the basis of expected grade level reading skills.

The statewide Florida Comprehensive Assessment Test (FCAT) was administered by school personnel in March 2005 to the third, fourth, and fifth grade subjects in the proposed study. The FCAT provides two measures for reading. One is based upon the

Florida Sunshine State Standards (SSS) and provides Developmental Scores that allow assessment of reading progress from year to year. The FCAT SSS also provides reading levels. Level 1 is considered failure while Levels 2 through 5 are considered passing. Level 3 is considered to represent grade level reading. The second part of FCAT assessment is a norm referenced test (NRT). The NRT compares Florida's students in reading and mathematics to other students within the U.S. and yields percentile rankings. The FCAT SSS Developmental Scores were used in this study as the measure for final analysis. According to the Florida Department of Education, the FCAT is a reliable and valid test and meets or exceeds standards for technically sound achievement tests (Assessment and Accountability Briefing Book, 2004).

The Collective Teacher Efficacy Scale – Short Form (CTES) was developed by Goddard (2002) and consists of 12 items using a Likert-type 6-point scale. Response choices range from *strongly disagree* to *strongly agree*. A copy of the CTES is provided in Appendix A. A significant positive correlation ( $r = .983$ ) was established between the short form and the original 21-item version. Using a multilevel predictive validity model, Goddard found that the scale was a positive predictor of variability in schools' mathematics achievement. Both the original scale developed by Goddard, Hoy, and Hoy (2000) and the short form, (Goddard), have been validated and determined to have strong reliability. Scoring of the CTES for this study required that six items from the scale (4, 5, 8, 10, 11, and 12) be reversed scored so that high scores would indicate high levels of collective efficacy.

## Data Collection

Permission was obtained from the school district's Research Oversight Committee to conduct this study. Further, the principals from Elementary A and B gave permission for their schools to participate and to release identification numbers of students who were limited English proficient. Copies of the principal permission forms are provided in Appendices C and D. Student identification numbers were then used to access test scores from the district's electronic databases.

As noted above, reading level data from the DRA and Flynt-Cooter were used to first establish reading levels of the two groups during the 2003-2004 school year and at the initial participation of students in the Reading First program. Test data from the GRADE administered in May 2004 and May 2005 were obtained from the school district's electronic database *Data Warehouse*. The two administrations of the GRADE served as a pre- and post-test for gauging reading achievement. The GRADE was scored by the school district and results were posted in *Data Warehouse* by the end of May 2005.

All subjects had an evaluation of their English language proficiency using the BVAT prior to being selected for participation in the study. Results from the BVAT are in the school district's database *Terms*, and were accessed by student identification numbers.

The DIBELS assessment is an integral and state required part of the Reading First program. All 34 Elementary A participants were administered the DIBELS during May 2005. For students in grade 2, scores were available and analyzed on the DIBELS

Nonsense Word Fluency and Oral Reading Fluency tests. For students in grades 3-5, only the Oral Reading Fluency test was administered.

To assess collective teacher efficacy in each school, a packet for each teacher was created and delivered to both Elementary A and B. Both principals requested that they be responsible for distributing the packets and returning them to the researcher. Each packet contained a Teacher Questionnaire (Appendix B), the collective teacher efficacy measure (CTES) (Appendix A), two Teacher Consent forms (Appendix E), and an envelope. All packets were delivered two weeks before the end of the teachers' contract. The teachers returned their sealed envelopes to the each principal who then provided them to the researcher at the end of the final week of school in May 2005. Teachers at Elementary A and B at all six grade levels and the Exceptional Student Education programs completed the CTES and Teacher Questionnaire except for third grade teachers at Elementary B. By the end of May 2005, no CTES or questionnaires had been returned by this group of teachers.

### Variables

The independent variable for this study was participation in a Reading First school. This variable could not be manipulated by the researcher as the school district had already established which schools were participants in Reading First. The majority of teachers at Elementary A were trained in Reading First methodology. The dependent variable was reading achievement as measured by standardized tests. Additionally, assessment of collective teacher efficacy served as a second dependent variable.

Subject selection variables for both groups included those students who met the following criteria: (a) were limited English proficient, (b) participated in ESOL during the 2004-2005 school year, (c) had no prior grade retentions, (d) were not identified as learning disabled, language impaired, or emotionally handicapped (e) had been tested with the GRADE during May 2004 and May 2005, (e) were students in grades 2 through 5 in either Elementary A or B, (f) first language was Spanish, and (g) participated in the district's free or reduced cost lunch program.

### Data Analysis

All data were entered into a research database using the *Statistical Package for the Social Sciences: Graduate Pack 11.5 for Windows* (SPSS, 2002) for the purpose of analysis. Data were grouped according to the questions of this study. Significance for all statistical analyses was determined at the .05 level.

The first research question of the study addressed the reading achievement levels of the two groups of students after the final administration of the GRADE in May 2005. Descriptive statistics including means and standard deviations were calculated for each group on the Total GRADE test and the individual subtests.

The second research question asked if a significant difference was seen between the two groups in the amount of reading progress made during the 2004-05 school year. To address this question, two-way analyses of variance (ANOVA) and repeated measures ANOVA were calculated on the basis of GRADE scores from the 2004 and 2005

administrations. Between-subjects factors included group participation, grade levels, and comparison of grade level combinations including 2-3 and 4-5.

For the third research question, it was asked what the correlation was between the DIBELS scores and the GRADE scores for the Reading First group. Pearson product-moment correlation was used to address this question.

The fourth research question also pertained to correlation and specifically asked if there was a relationship between the Total Score on the GRADE and the FCAT scores for the third, fourth, and fifth grade students in both groups. Again, Pearson product-moment correlation was used to address this question.

For research question five, the researcher asked if the GRADE and DIBELS scores could serve as predictors of the FCAT scores for students in grades 3-5. Linear regression analyses were performed for the GRADE/FCAT and DIBELS/FCAT combinations. Additionally, multiple regression analysis was performed with both the DIBELS and GRADE scores to answer this question.

The final research question asked if there was a significant difference in collective teacher efficacy between the Reading First school and the non-Reading First school based upon the CTES. The teachers in each school were considered as two independent groups and to address this question, a t-test for independent means was calculated.

## CHAPTER FOUR

### ANALYSIS OF DATA

#### Results as Related to Research Questions

A database was developed that included data from 34 students at Elementary A, the Reading First school, and 34 from Elementary B, non-Reading First, in grades 2-5. Additionally, scores obtained from the Collective Teacher Efficacy Scale (CTES) completed by teachers at both schools were also included. Students were eliminated from Elementary A who had not attended that school for the entire 2004-2005 school year. Test scores were collected from the school district's *Data Warehouse* electronic database and were analyzed using *Statistical Package for the Social Sciences: Graduate Pack 11.5 for Windows* (SPSS, 2002).

To establish the initial reading levels for both groups of students, data were obtained from the 2003-04 academic year as described above. Table 4 shows results based upon an informal reading inventory (DRA) for grades 2-5. As can be seen, 20 of the 25 students (80%) with available DRA scores at Elementary A were reading at a second grade level (DRA 20 -24) or lower. At Elementary B, all 27 (100%) students were reading at this level. The remaining 5 students (20%) at Elementary A were at a fourth to fifth grade level. These results indicated that the majority of students at each school were at a low and similar reading level, particularly at second and third grades where all students were reading at a first grade level or lower.

## Question One

What are the reading achievement levels of a Reading First group (Elementary A) and a non-Reading First group (Elementary B) of ESL students after a posttest on the Total Test and the individual subtests of the Group Reading Assessment and Diagnostic Evaluation (GRADE) (Williams, 2002)?

Descriptive statistics involving means, ranges, and standard deviations were used to analyze the Total Test scores from the GRADE. Both schools had data available for a total of 34 students each who met selection criteria. The GRADE Standard Score was derived on the basis of a national normative sample and has a mean of 100 and a standard deviation of 15. At Elementary A, the mean Total Score was 90.06 ( $SD = 9.90$ ) while at Elementary B, the mean was 94.24 ( $SD = 10.77$ ) for the 2005 administration. The difference between the two means was less than one  $SD$  (15 points) from each other and both were within one  $SD$  of the GRADE's mean of 100. The range for both sets of data was 43.

Means and standard deviations were also calculated for grades 2 and 3, and 4 and 5 at each school for the 2005 administration. While Elementary A emphasized Reading First strategies at all grade levels, grades 2 and 3 were the primary targeted recipients of Reading First resources. When GRADE scores from second and third grades at both schools were analyzed, a difference of one standard deviation or greater was not seen between the two means. Elementary A's mean for the combination of grades 2 and 3 was 90.56 ( $SD = 11.00$ ) ( $N = 16$ ), while Elementary B's mean was 92.61 ( $SD = 12.37$ ) ( $N = 18$ ). The range of scores for both schools was 43 for grades 2 and 3.

When scores at the fourth and fifth grade levels were combined at both schools, the mean for Elementary A was 89.61 ( $SD = 9.10$ ) ( $n = 18$ ) while at Elementary B, the mean was 96.06 ( $SD = 8.66$ ) ( $n = 16$ ). Again, a significant difference (15 points or greater) was not seen. The ranges of scores at grades 4 and 5 for Elementary A and Elementary B were 39 and 33, respectively. Table 5 provides a summary of the Total GRADE score descriptive statistics.

In comparing the 2004 and 2005 Total GRADE scores for both schools, it was noted for Elementary A that 15 students (44%) had lower scores during the 2005 administration while 19 (56%) had higher scores. For Elementary B, 7 (21%) students had lower scores, 26 (76%) had higher scores, and one student had no change in score.

Table 5

## Elementary A and B Descriptive Statistics for Total Grade Standard Scores

	<i>N</i>	Minimum	Maximum	<i>M</i>	<i>SD</i>	Range
Elementary A						
Total GRADE	34	68	111	90.06	9.90	43
Grades 2-3 Total Score	16	68	111	90.56	11.00	43
Grades 4-5 Total Score	18	71	110	89.61	9.11	39
Elementary B						
Total Grade	34	71	114	94.24	10.77	43
Grades 2-3 Total Score	18	71	114	92.61	12.37	43
Grades 4-5 Total Score	16	78	111	96.06	8.66	33

Note: All data are from the 2005 administration of the GRADE.

For each individual subtest from the GRADE, the electronic database reported the number of items attempted, the number answered correctly, the total possible number of items, and the percentage of items answered correctly. Scale scores or any other score types were not reported. To examine each student's performance on the individual subtests, the number of items answered correctly was analyzed and descriptive statistics were applied for each grade level. An analysis for all grades combined at each school was not possible since the specific subtests administered were dependent on the students' grade level. At grade 2, the Listening Comprehension, Passage Comprehension, Sentence Comprehension, Word Meaning, and Word Reading subtests were given. At grade 3, Listening Comprehension, Passage Comprehension, Sentence Comprehension,

Vocabulary, and Word Reading tests were administered. For grades 4 and 5, the Listening Comprehension, Passage Comprehension, Sentence Comprehension, and Vocabulary subtests were given. While subtests can have the same title, the total possible number of items can vary by grade level. Table 6 provides the means for each school on the subtest scores (number answered correctly) at each grade. Following each rounded mean, the total numbers of items possible with the percentage correct are provided below in parenthesis. What emerged as particularly salient from these data were the relatively low percentages for the Vocabulary subtest. At the fourth and fifth grade levels, this was the lowest percentage except for both schools. Also, the consistently highest percentages were seen on the Word Meaning and Word Reading subtests for both schools.

Table 6

Mean Number of Items and Percentages Answered Correctly on GRADE Subtests

	LC <sup>a</sup>	PC	SC	VOC	WM	WR
Elementary A						
Grade 2	14/17 <sup>b</sup> (82) <sup>c</sup>	17/28 (61)	11/19 (58)	----	26/27 (96)	27/28 (96)
Grade 3	12/17 (71)	11/28 (39)	12/19 (63)	14/30 (47)	----	27/30 (90)
Grade 4	13/17 (76)	15/28 (54)	12/19 (63)	15/35 (43)	----	----
Grade 5	14/17 (82)	18/30 (60)	10/19 (53)	11/35 (31)	----	----
Elementary B						
Grade 2	15/17 (88)	16/28 (57)	14/19 (74)	----	25/27 (93)	25/28 (89)
Grade 3	14/17 (82)	15/28 (54)	14/19 (74)	17/30 (57)	----	28/30 (93)
Grade 4	15/17 (88)	19/28 (68)	15/19 (79)	21/35 (60)	----	----
Grade 5	15/17 (88)	16/30 (53)	12/19 (63)	14/35 (40)	----	----

<sup>a</sup> LC = Listening Comprehension, PC = Passage Comprehension, SC = Sentence Comprehension  
 VOC = Vocabulary, WM = Word Meaning, WR = Word Reading

<sup>b</sup> Ratio of number answered correctly to total possible items for each subtest

<sup>c</sup> Percentage of number answered correctly out of total possible

Blank spaces indicate that the subtest was not given at that particular grade level.

## Question Two

Is there a significant difference between the two groups in the amount of progress made during the 2004-2005 school year as shown by the GRADE?

Analyses were conducted to examine the interaction of grade placement, pre and posttest GRADE scores with participation in Reading First and non-Reading First groups. Separate two-way ANOVAs were computed, first with the pretest Total GRADE scores and second with the posttest scores. For the pretest analysis, group participation (Reading First or non-Reading First) and grade level (grades 2-5) served as fixed factors with the GRADE scores serving as the dependent variable. There was not a statistically significant effect of grade level ( $F_{3,60} = 0.73, p > .05$ ). Grade level accounted for approximately 4% of the variance in score. The means for each grade level were: Second grade 92.06, third grade 88.17, fourth grade 89.25, and fifth grade 86.93. There was not a statistically significant effect for group participation ( $F_{1,60} = 0.19, p > .05$ ). The Reading First group's pretest mean of 88.43 was not significantly different from the non-Reading First mean of 89.62. Less than 1% of the variance in pretest score could be attributed to group. Finally, there was not a statistically significant interaction effect ( $F_{3,60} = 1.55, p > .05$ ). The interaction between grade and group accounted for approximately 7% of the variance in score. Table 7 provides a summary of the pretest analyses.

For the posttest GRADE Total scores, similar results were found. There were no statistically significant findings for grade level ( $F_{3,60} = 2.03, p > .05$ ), group effects ( $F_{1,60} = 2.98, p > .05$ ), or interaction between grade level and group ( $F_{3,60} = 0.91, p > .05$ ). Grade level accounted for approximately 9% of the variance while group accounted for approximately 5%. The interaction of grade level and group accounted for approximately

4% of the variance. Table 7 provides specific data from analyses of the posttest measure. Means for posttest GRADE scores were: Second grade 95.69, third grade 88.06, fourth grade 92.85, and fifth grade 92.36. For the two groups, there was no statistical difference between Reading First ( $M = 90.30$ ) and non-Reading First ( $M = 94.70$ ) mean scores.

To further address research question two, a repeated measures one-way analysis of variance (ANOVA) was computed with pre and post Total GRADE test scores serving as the within-subjects factors (Test) and the four grade levels (2-5) as the between-subjects factors (Group). There was a statistically significant difference seen between pretest ( $M = 89.15$ ,  $SD = 11.04$ ) and posttest ( $M = 92.15$ ,  $SD = 10.48$ ) scores ( $F_{1, 64} = 7.44$ ,  $p < .01$ ). Approximately 10% of the variance in scores could be attributed to the pre/post factor. There was not a statistically significant difference seen for grade levels ( $F_{1, 64} = 1.07$ ,  $p > .05$ ). Means for each grade were: Second grade 93.88 ( $SD = 2.42$ ), third grade 88.11 ( $SD = 2.28$ ), fourth grade 91.05 ( $SD = 2.16$ ), and fifth grade 89.64 ( $SD = 2.58$ ). Approximately 5% of the variance could be explained by between-subjects effects. Finally, there was not a statistically significant interaction effect between groups and GRADE pre and post scores ( $F_{3, 64} = 1.02$ ,  $p > .05$ ). Table 8 provides a summary of all repeated measures ANOVA results.

Additional repeated measures ANOVAS were conducted to specifically examine combined data from both schools where grades 2 and 3 and grades 4 and 5 served as two separate groups. As noted above, students in the second and third grades at Elementary A were targeted participants in the Reading First program. A one-way repeated measures ANOVA was computed with the second and third grade pre and post Total GRADE scores serving as within-subjects factors (Test) and the two grade levels (2 and 3) serving

as between-subjects factors Group). A statistically significant difference was not seen in pre ( $M = 90.0$ ,  $SD = 12.23$ ) and posttest ( $M = 91.65$ ,  $SD = 11.62$ ) scores for the two groups ( $F_{1,32} = .72$ ,  $p > .05$ ). Approximately 2% of the variance could be explained by the pre and post test factors. When considering group participation no statistical difference seen between grade 2 ( $M = 90.44$ ,  $SD = 2.70$ ) and grade 3 ( $M = 91.17$ ,  $SD = 2.55$ ) groups ( $F_{1,32} = .04$ ,  $p > .05$ ). Less than 1% of the variance could be attributed to group participation. Similarly, a significant interaction effect was not seen ( $F_{1,32} = .51$ ,  $p > .05$ ). Approximately 2% of the variance could be explained by the interaction of groups and test administration.

In contrast to the second and third grade results, a statistically significant difference was seen between pre ( $M = 88.29$ ,  $SD = 9.8$ ) and posttest ( $M = 92.65$ ,  $SD = 9.36$ ) scores for grades 4 and 5 ( $F_{1,32} = 10.71$ ,  $p < .01$ ). Approximately 25% of the variance in scores could be attributed to pre/posttest factor. When group participation was considered, there was no statistically significant difference seen between grade 4 ( $M = 91.05$ ,  $SD = 1.98$ ) and grade 5 ( $M = 89.64$ ,  $SD = 2.34$ ) students ( $F_{1,32} = .21$ ,  $p > .05$ ). Approximately 7% of the variance could be accounted for by group participation. A statistically significant interaction effect was not found ( $F_{1,32} = .44$ ,  $p > .05$ ). Here, approximately 1% of the variance could be explained by the interaction of group participation and the pre/post testing.

Table 7

Two-Way ANOVA for GRADE Pre and Posttest Scores, Grades 2-5, Group, and Interaction Effects

Condition	Variable	F	df	$\eta^2$
Pretest (n = 68)				
	Grade	0.73	3, 60	.035
	Group	0.19	1, 60	.003
	Grade X Group	1.55	3, 60	.072
Posttest (n = 68)				
	Grade	2.03	3, 60	.092
	Group	2.98	1, 60	.047
	Grade X Group	0.91	3, 60	.044

Note: \*  $p \leq .05$ , \*\*  $p \leq .01$

Grade = Grades 2-5, Group = Reading First and Non-Reading First

Table 8

## Repeated Measures ANOVA

Condition	Variable	F	df	$\eta^2$
Grades 2-5 (n= 68)				
	Group	1.07	3, 64	.048
	Test	7.44**	1, 64	.104
	Group X Test	1.02	3, 64	.046
Grades 2 and 3 (n =34)				
	Group	0.04	1, 32	.001
	Test	0.72	1, 32	.022
	Group X Test	0.51	1, 32	.016
Grades 4 and 5 (n =34)				
	Group	.21	1, 32	.066
	Test	10.71**	1, 32	.251
	Grade X Group	.44	1, 32	.014

Note: \*  $p \leq .05$ , \*\* $p \leq .01$

Group = Grade Levels, Test = Pre and Post GRADE Scores

## Question Three

What is the correlation between scores on the Dynamic Indicators of Basic Early Literacy Skills (DIBELS), (Good & Kaminski, 2001) and GRADE scores for the Reading First group?

To address this question, two sets of data were derived and were then analyzed by computing Pearson product-moment correlations. First, all students in the Reading First group had been administered the Oral Reading Fluency (ORF) subtest from the DIBELS. The number of words read correctly in one minute was reported as the raw score. These scores were then correlated with the Total GRADE scores from the 2005 administration.

A significant positive correlation was seen ( $r = .38, p < .05$ ) between the GRADE and the DIBELS ORF scores for all 34 students. Approximately 14% of the variance could be accounted for.

A second correlation was computed between the Total GRADE and the Nonsense Word Fluency (NWF) subtest scores for the students in second grade ( $N = 8$ ). This was the only sub-group from the Reading First sample tested with the NWF subtest. Here, the final score was represented by the total number of letter sounds students could provide for nonsense words within one minute. Results showed a non-significant negative correlation between the NWF subtest and the total GRADE score for this small sample ( $r = -.08, p > .05$ ). However, meaningful interpretation is limited due to the small sample size. Table 9 provides a summary of correlation coefficients.

#### Question Four

What is the correlation between Total Score on the GRADE and the Florida Comprehensive Assessment Test (FCAT) for third, fourth, and fifth grade students in the Reading First and non-Reading First groups?

For the Reading First group, Pearson product moment correlations were computed for all students in grades 3-5 who had available FCAT scores ( $N = 25$ ). A significant positive correlation was found between the FCAT Developmental Scores and the Total GRADE scores for this group ( $r = .56, p < .01$ ). Here, approximately 31% of the variance could be explained. Interestingly, when the Reading First 2004 Total GRADE scores

were correlated with the 2005 FCAT Developmental scores, a significant correlation was not found ( $r = .25, p > .05$ ).

Correlations were computed for each grade level with both groups using 2005 GRADE Total and 2005 FCAT Developmental Scores. Among the Reading First group, only the fourth grade group ( $n = 13$ ) showed a significant positive correlation ( $r = .61, p < .05$ ). From this calculation, approximately 37% of the variance could be explained. Correlations for grade 3 ( $n = 7$ ) ( $r = .55, p > .05$ ) and grade 5 ( $n = 5$ ) ( $r = .60, p > .05$ ) were not statistically significant using a 2-tailed test.

For the non-Reading First group ( $N = 25$ ), a significant positive correlation was also seen between the 2005 FCAT Developmental Scores and the GRADE Total scores ( $r = .80, p < .01$ ) for students in grades 3-5. From this computation, 64% of the variance could be explained. Similar to the Reading First group, a non-significant correlation was seen between the 2005 FCAT scores and the 2004 Total GRADE scores ( $r = .23, p > .05$ ).

When correlations were performed on individual grade level FCAT and GRADE scores in the non-Reading First group, all three calculations were statistically significant. At grade 3 ( $n = 9$ ) results yielded a Pearson product moment  $r$  of .90 ( $p < .01$ ). Here, 81% of the variance could be accounted for. For grade 4 ( $n = 7$ ), the correlation coefficient was .86 ( $p < .05$ ) and could explain approximately 74% of the variance. For grade 5 ( $n = 9$ ) of the non-Reading First group, results revealed a correlation coefficient of .70 ( $p < .05$ ) and could account for 49% of the variance. It is important to note that interpretation of individual grade level correlations is very limited due to the small number of scores in

each group. Table 9 provides a summary of correlation coefficients for variables addressed in research question 4.

Table 9  
DIBELS, FCAT, and GRADE Correlations

Variables	GRADE (2005)	GRADE (2004)
Reading First		
DIBELS ORF	.38*	---
DIBELS NWF	-.08	---
FCAT DS	.56**	.25
3 <sup>rd</sup> Grade FCAT	.55	---
4 <sup>th</sup> Grade FCAT	.61*	---
5 <sup>th</sup> Grade FCAT	.60	---
Non-Reading First		
FCAT DS	.80**	.23
3 <sup>rd</sup> Grade FCAT	.90**	---
4 <sup>th</sup> Grade FCAT	.86*	---
5 <sup>th</sup> Grade FCAT	.70*	---

Note: \*  $p \leq .05$ , \*\*  $p \leq .01$

#### Question Five

Can the GRADE and/or DIBELS scores serve as predictors of the FCAT Reading score?

The predictive capability of the Total GRADE score and the DIBELS Oral Reading Fluency (ORF) for the Reading First group were addressed individually through regression analysis. The Total GRADE score for the third, fourth, and fifth grade students ( $N = 25$ ) was found to be a significant predictor of the 2005 FCAT Developmental Score ( $r = .56$ ,  $F_{1,24} = 10.67$ ,  $p < .01$ ). Because  $r^2 = 0.29$ , 29% of the variance in the FCAT

score could be accounted for by the GRADE score. The regression equation was found to be: Predicted FCAT Score =  $-260.39 + 16.89$  (GRADE Total score).

The DIBELS ORF (N = 25) score was also found to be a significant predictor of the FCAT score for the Reading First group ( $r = .44$ ,  $F_{1,24} = 5.53$ ,  $p < .05$ ). Because the  $r^2$  value was .16, 16% of the variance in FCAT Reading score could be accounted for by the ORF score. The regression equation was found to be: Predicted FCAT score =  $784.21 + 4.4$  (ORF score).

A multiple regression was performed with the FCAT Reading Developmental Score as the dependent variable and the DIBELS ORF and GRADE Total scores as the combined predictor variable for the Reading First group. The combination of the ORF and GRADE scores was found to be a significant predictor of FCAT scores ( $R = .62$ ,  $F_{2,24} = 6.82$ ,  $p < .05$ ). This combination could account for 33% ( $R^2 = .33$ ) of the variance in the FCAT score. The regression equation was: Predicted FCAT score =  $-286.28 + 2.75$  (ORF score) +  $13.96$  (GRADE Total score).

For the non-Reading First group (N = 25), a regression analysis was performed using the FCAT Developmental Score as the dependent variable and the Total score from the GRADE as the predictor variable. Similar to the Reading First group, the GRADE Total score was a significant predictor of the FCAT score ( $r = .80$ ,  $F_{1,24} = 39.63$ ,  $p < .01$ ). The  $r^2$  value was .62 and as a result, 62% of the variance of the FCAT score could be accounted for by the Total GRADE score. The regression equation was found to be: Predicted FCAT score =  $-1725.23 + 32.22$  (GRADE Total score).

Results of the regression analyses of the FCAT Developmental Score for both groups on the predictor variables described above are outlined in Table 10.

Table 10

Predictors of FCAT Developmental Scores for Reading First and Non-Reading First Groups

Predictor Variable	<i>MS</i>	<i>df</i>	<i>F</i>	$r^2/R^2$
Reading First Group				
GRADE Total	43786.13	1, 24	10.67**	.29
DIBELS ORF	51666.72	1, 24	5.53*	.16
GRADE & ORF	41361.13	2, 24	6.82*	.33
Non-Reading First Group				
GRADE Total	52139.97	1, 24	39.63**	.62

Note: \*  $p \leq .05$ , \*\*  $p \leq .01$

### Question Six

Is there a significant difference in collective teacher efficacy, as assessed by the Collective Teacher Efficacy Scale (CTES) (Goddard, 2002), between the Reading First and the non-Reading First school?

First, the CTES was analyzed to determine its reliability. Items would be deleted that resulted in a negative item-total correlation or a correlation of .15 or lower to the total scale. This would help reduce variables such as poor item construction, random response, or others related to measurement error. Tables 11 and 12 provide item statistics for the Reading First and Non-Reading First groups, respectively from the 12-item CTES. Table 13 provides statistics for both groups combined. As can be seen, no items met

criteria for deletion. Responses were on a 6-point Likert-type scale with high scores indicating higher levels of collective efficacy. The following items were reversed scored: 4, 5, 8, 10, 11, and 12.

An estimate of internal consistency was first derived by computing Cronbach's coefficient alpha (Cronbach, 1951) on each group's scale individually. Results for the Reading First group indicated evidence of internal consistency with an alpha coefficient of .85. For the non-Reading First group, an alpha coefficient of .70 was obtained. Second, the two groups were combined so that alpha coefficients could be obtained for each item and all 68 total scores. Alpha levels ranged from .72 to .84 for the individual items and the total scores.

Table 11

Item Statistics for the Reading First CTES

Item No.	M	SD	Item-Total <i>r</i>
1	4.58	1.11	.58
2	4.61	.79	.78
3	5.16	1.00	.65
4	5.13	1.21	.69
5	5.66	.63	.43
6	2.97	1.35	.73
7	1.66	.78	.63
8	3.79	1.44	.64
9	2.53	1.20	.68
10	4.90	1.16	.55
11	4.21	1.38	.53
12	4.71	1.39	.60
Total	49.89	8.41	--

Table 12

Item Statistics for the Non-Reading First CTES

Item No.	M	SD	Item-Total <i>r</i>
1	5.07	.69	.47
2	5.23	.68	.47
3	5.47	.57	.18
4	5.23	1.36	.40
5	5.47	.86	.33
6	3.60	1.50	.48
7	2.90	1.54	.46
8	4.37	1.33	.73
9	4.17	1.23	.51
10	5.43	1.25	.39
11	5.33	1.06	.47
12	5.27	.64	.38
Total	57.23	6.75	--

Table 13

## Item Statistics for Both the Reading First and Non-Reading First CTES

Item No.	M	SD	Item-Total <i>r</i>
1	4.79	.97	.58
2	4.88	.80	.72
3	5.29	.85	.54
4	5.18	1.27	.56
5	5.56	.76	.31
6	3.25	1.44	.66
7	2.16	1.28	.67
8	4.07	1.40	.70
9	3.19	1.47	.72
10	5.22	1.12	.56
11	4.68	1.38	.62
12	4.94	1.16	.55
Total	53.13	8.50	--

A *t*-test for independent means was computed using the total scores from the Reading First and non-Reading First groups' CTES. As noted above, six items were reversed scored so that the higher the total scores, the greater the indication of collective efficacy. Results revealed a statistically significant difference in collective teacher efficacy between the Reading First and non-Reading First school ( $t = -3.89$ ,  $df = 66$ ,  $p < .01$ ). The 95% Confidence Interval indicated the true mean difference (-7.34) may range from  $-11.11 < \mu < -3.57$ . Teachers completing the CTES in the Reading First school had a mean score of 49.89 ( $SD = 8.42$ ) while those in the non-Reading First school had a significantly higher mean of 57.23 ( $SD = 6.76$ ).

CHAPTER FIVE  
CONCLUSIONS, DISCUSSION, AND RECOMMENDATIONS

Summary of Findings

The focus of this study was the reading achievement of Hispanic, limited English students in a school designated as Reading First and one that was not a recipient of Reading First resources. Reading achievement was assessed with the Group Reading Assessment and Diagnostic Evaluation (GRADE), (Williams, 2002), a standardized, group administered, and nationally normed reading test. Contrasts between pre and post administrations of the GRADE were examined and were considered in relation to Reading First and non-Reading First groups as well as, individual grade levels. Other measures of reading skills that were available at the time of the study were also utilized as dependent measures in assessing Reading First and non-Reading First participants. These included the Florida Comprehensive Assessment Test (FCAT) and the Dynamic Indicators of Basic Early Literacy Skills (DIBELS), (Good & Kaminski, 2001). As a secondary investigation, the concept of collective teacher efficacy was assessed in an effort to determine if differences were apparent between the Reading First and non-Reading First schools. One hypothesis considered was that the intensive teacher training and focus of Reading First strategies in a school community could facilitate increased collective teacher efficacy.

Results from an informal reading inventory were considered for initial assessment of reading skills at the time the Reading First program was initiated at the beginning of

the 2003-2004 academic year. Essentially, reading levels for both groups were low and for the majority, at or below first grade. At Elementary B, the non-Reading First school, 100% of students were at this level. For the Reading First school, five of the 34 students had informal reading scores at a fourth to fifth grade level while the remaining 80% were at first grade or lower.

The GRADE served as a pretest for both schools during May 2004 and was administered again to all participants during May of 2005. When compared to the national normative sample, the pretest and posttest Total Grade scores for both groups were within one standard deviation (*SD*) (15 points) of each other and within one *SD* of the test's mean Scale score of 100. When the individual subtests of the post GRADE were examined, findings suggested that both groups performed well on Word Reading and Word Meaning but had rather consistent difficulty with the Vocabulary measure. This would not be surprising since all students in this study were limited English proficient.

To assess the potential significance of any differences between the two groups and the pre and posttest measures, two-way and repeated measures analyses of variance (ANOVA) were computed. Findings from the two-way ANOVAs did not reveal statistically significant differences between the Reading First and non-Reading First groups on either the pretest or posttest measures of the GRADE. Further, significant effects for grade level or the interaction of group participation with grade level were not found for either the pretest or posttest analyses. However, when considering the two groups as a whole on the pre and posttest Total grade scores, a significant difference was seen based upon repeated measures ANOVA. The posttest mean score was significantly

higher than the pretest mean. This indicated that the students as a whole had progressed significantly since the pretest in their reading skills as assessed by the GRADE.

Statistically significant interaction effects between the two groups and the GRADE scores were not found.

While all students at Elementary A were recipients of Reading First strategies, second and third graders represented the primary targeted group for Reading First resources. When GRADE data from second and third grade were analyzed as an individual group at both schools on the pre and post Total GRADE scores, no statistical difference was found. Conversely, when fourth and fifth grades were isolated, a statistically significant difference was found between pre and posttest measures. Posttest scores were significantly higher than pretest. Significant differences between the Reading First and non-Reading First groups were not found nor was there a significant interaction effect. Essentially, this grouping of students made substantial gains from the 2004 to the 2005 academic year. Further, this group appeared to have contributed significantly to the overall difference found in pre and posttest measures as noted above.

Correlational analyses were used to address questions about the relationship among the various tests instruments used in this study. All students in the Reading First group were administered the DIBELS four times during the 2004-05 academic year. Scores from the last administration during May 2005 were correlated with the Total post-GRADE scores. A significant positive correlation was found between the DIBELS Oral Reading Fluency (ORF) test and the Total GRADE. Similarly, a significant and positive correlation was found between the Total post-GRADE scores and Developmental Scores from the FCAT reading measure for both groups. When individual correlations were

computed for each grade level, only the fourth grade scores ( $n = 13$ ) were positively related to the FCAT for the Reading First group. Conversely, scores from all non-Reading First grade levels (2-3) were found to have significant and positive correlations with the FCAT Developmental Scores. While correlational data at individual grades and the DIBELS NWF and GRADE relationship were presented, it was noted that meaningful interpretation was limited due to very small sample sizes. When the 2004 GRADE scores from both groups were correlated with the FCAT scores, a statistically significant relationship was not found.

Analyses were also conducted to address whether the Total GRADE scores and the DIBELS scores could serve as predictors of the FCAT. Results showed that the GRADE Total score was a statistically significant predictor of the FCAT reading Developmental Score for both groups of students. The DIBELS ORF test was also found to be a significant predictor of the FCAT reading score for the Reading First group. A significant correlation was not found between the DIBELS Nonsense Word Fluency and FCAT Developmental Scores.

As indicated above, students in the non-Reading First group were not given the DIBELS. Further, when the DIBELS ORF and the Total GRADE scores were combined, the two measures together accounted for greater variance in the FCAT scores than either alone.

As a second part of this study, the concept of collective teacher efficacy was assessed by administering the Collective Teacher Efficacy Scale-Short Form (CTES) (Goddard, 2002) to teachers at both Elementary A and B. After distribution of the CTES at each school, the return rate was 69.8% at Elementary A and 60% at Elementary B.

Unfortunately, none of the third grade teachers at Elementary B returned the CTES and as a group, their input was not represented in the overall findings. Analysis of the returned CTES indicated significant internal consistency for the CTES from both schools (Elementary A,  $\alpha = .70$ ; Elementary B,  $\alpha = .85$ ). None of the 12 items required deletion due to negative item-total correlation or a correlation of .15 or less. Findings revealed a significantly higher level of collective teacher efficacy at Elementary B, the non-Reading First school, than at Elementary A.

## Discussion

### Reading First and ESL Students

The literature on explicit reading instruction for ESL students reviewed as a part of this study was sparse. However, there was tentative evidence that these students could benefit from such instruction, particularly when phonemic awareness training was included (Gunn et al., 2000, 2002). The present study focused on the benefits of systematic and explicit reading instruction for limited English proficient students with specific attention given to the instructional reading strategies that are provided in the Reading First program. Findings did not conflict with the literature. After participating a year in a Reading First school where resources and instructional methods were focused on phonemic awareness, phonics, fluency, vocabulary, and comprehension, a sample of ESL students did not score significantly different from a sample of ESL students instructed in a non-Reading First school on a group administered reading test. Both

groups of students had similar outcomes after one year of instruction regardless of Reading First emphasis or not. A significant difference was seen between pre and posttest measures on the group test when both samples of students were combined. At first glance, both Reading First and non-Reading First groups apparently benefited from the reading instruction available at their schools. Further analysis suggested that this difference in pre and posttest scores was partially due to the progress made by the fourth and fifth graders. When second and third grade students from both schools were combined, no difference was found between the pre and posttest measures. It would appear that both schools provided similar reading instruction. Another interpretation is that as the level of English proficiency increased for students during the year and particularly for fourth and fifth graders, they were better able to benefit from reading instruction provided only in English. In fact, there were students in both groups who were exited from the ESOL programs by the end of the year.

When the GRADE reading data collected here were compared to a national normative sample, both Reading First and non-Reading First groups were within one standard deviation of the test's mean of 100 on both the pretest and posttest. An examination of the individual subtests comprising the GRADE indicated higher scores on tasks involving word identification and lower skills in vocabulary. The latter was found primarily for the fourth and fifth grade students at both schools. It would appear that limited English proficient students benefited from instruction in phonemic awareness and phonics and in turn, these skills likely helped their performance on the Word Reading and Word Meaning subtests from the GRADE. Consistent with their limited English status, it is understandable that vocabulary knowledge would be lower. It would also be expected

that limited English ability influenced measures of reading comprehension where a higher level of abstract language might have been required. However, this was an inconsistent finding from the GRADE data collected for this study. Generally, interpretation of findings suggested that ESL students benefited from instruction emphasizing explicit reading instruction. They did not, in fact, appear to lose skills or fail to make reading progress over the course of one year. However, what remained unanswered was the ultimate value of the extensive resources and expense of Reading First for ESL students and whether their instructional time may have been better spent in a dual language program or other options where native language instruction occurred. Tentative findings here suggested that continued and intensive efforts will be needed to address vocabulary development and the students' cognitive academic language skills. At least in Florida, expectations for adequate language proficiency and grade level reading achievement appear to arise prior to the time when such development normally occurs for students in immersion programs (Collier, 1989).

Correlation analyses showed a positive relationship among three dependent measures used to assess reading achievement. Using multiple regression, two of the measures, GRADE Total Scale score and the DIBELS Oral Reading Fluency (ORF) measure, were predictive of the state FCAT reading Developmental Score and accounted for a greater amount of variance than either measure alone. This finding was particularly interesting since ORF assessed the number of words read correctly within one minute. The GRADE does not have a similar test but findings here suggested that some of the same processes such as decoding involved in reading words accurately aloud likely

contributed to the GRADE and FCAT measures. The GRADE's Word Meaning and Word Reading subtests appeared to rely heavily on decoding words.

The regression analyses tentatively suggested that the GRADE and DIBELS scores could serve to alert school and district personnel about ESL students who could potentially fail the FCAT. However, to actually benefit the most from the capabilities of the predictive measures, it would be essential that the district administer the GRADE prior to the FCAT. This was not the case for the present study.

The correlation between the 2004 administration of the GRADE and the 2005 FCAT scores was not significant for the students who participated in this study. One possible hypothesis was that as the students became more proficient in English fluency and reading, they were better able to generalize the skills they had acquired to variable measures of reading, especially if such measures became less context dependent.

As a final note, nearly half (46%) of students in this study who took the 2005 FCAT failed it based upon state criteria where Level 1 performance signified below grade level reading skills. Approximately 52% of Reading First and 40% of Non-Reading First students scored at Level 1. In view of the literature reviewed here and the theories pertaining to the amount of time required for second language acquisition, this finding would not appear extraordinary.

#### Collective Teacher Efficacy

Results here did not support the hypothesis that, at least in this study, a school's designation as Reading First would necessarily promote increased collective efficacy when compared to a non-Reading First school. The research question asked if the

intensive training of teachers in Reading First strategies and a school-wide approach to implementing them would contribute to positive teacher efficacy. As seen in Table 3, approximately 78% of teachers responding to the Teacher Questionnaire at Elementary A had received Reading First training. Results showed that collective efficacy was significantly higher in the non-Reading First (Elementary B) school. An obvious question was what were the variables that influenced this difference in collective efficacy? As noted by Ross, Hogaboam-Gray, and Gray in 2003, there was no available research that had examined principals' role in contributing to collective teacher efficacy. According to Ross et al., there were, however, studies supporting the influence of transformation leadership on collective efficacy. Leadership variables could not be excluded as contributing to the difference found in the present study.

Another variable that could be questioned as contributing to the difference may have been the degree of teacher experience and training. For the Reading First school, 24% of teachers were first year teachers while only 7% were first year at the non-Reading First school. For the latter, 64% of teachers responding to the Teacher Questionnaire had more than 5 years of teaching experience. At the Reading First school, 43% had more than 5 years. Years of college education also varied. At Reading First, 29% had degrees beyond the bachelor's level. In contrast, 50% of teachers responding to the questionnaire at the non-Reading First school had degrees beyond the bachelor's. At least in the area of student achievement, teacher experience and training have been found to have a significant effect on reading and mathematics achievement with somewhat smaller effects seen for training (Nye, Konstantopoulos, & Hedges, 2004). A final consideration would be the amount of time a staff has worked together and the degree of turnover. Could the

cohesiveness and mutual experience of a group positively influence collective efficacy? As noted above, Elementary B was a relatively new school and 82% of respondents had been there for a 2 to 5 year period. In contrast, approximately 32% of Elementary A respondents were first year teachers in that school.

### Conclusions

Collier's summary of research (1989) highlighted the challenges for ESL students in learning a second language in the U.S. while at the same time, acquiring academic skills. Her discussion clearly underlined the complexity of the issue and in turn, the necessity to avoid simplistic solutions and the failure to acquire evaluative data to assess their benefits. The present study offered no conclusive evidence that explicit reading instruction involving phonics, phonemic awareness, fluency, vocabulary, and comprehension, significantly helped or hindered the reading achievement of a sample of ESL students in a Southwest Florida school district. The group of Reading First students who met criteria for this study eventually attained a reading level equivalent to that of a sample who did not attend a Reading First school. Interpretation of findings suggested that both groups likely received similar instruction regardless of the Reading First designation or not and second, an increase in their English proficiency may have made English only reading instruction more meaningful. There is the implication as noted by August and Hakuta (1997), that the quality and nature of instruction provided by teachers is as important as the language of instruction. This may also be true regardless of programmatic instructional designations such as Reading First.

As a second part of the present study, the construct of collective teacher efficacy (CTES) was explored in an effort to understand its relationship to a school wide curricular approach such as Reading First. The outcome did not support the notion that Reading First necessarily and positively enhanced CTES. However, there was tentative and additional evidence that CTES is likely a complex and multifaceted concept with an array of variables that interact to help generate teachers' beliefs that they can positively influence the achievement of students in their school. From this study, it would seem worthwhile to continue the exploration of CTES and particularly, if there is indeed a mutual reinforcement of students' feelings of success and their teachers' collective efficacy beliefs. For ESL students, understanding how to enhance the positive interaction of CTES and student success could lend further support to understanding and addressing the challenges they face.

#### Recommendations for Future Research

This study represented a small step in exploring the benefits of explicit reading instruction for limited English, lower socioeconomic Hispanic students. It would seem paramount to continue the pursuit of appropriate reading instruction for such students particularly in the state of Florida where third graders can face multiple retentions if they fail the FCAT reading measures. It is unfortunate that the National Reading Panel (NRP) (National Institute of Child Health and Human Development, 2000a) chose not to consider research in their analyses pertaining to the reading instruction for ESL students. It is also unfortunate that programs such as Reading First have generalized the NRP's

findings to all students, including those where English is a second language (ESL) with, thus far, a poverty of empirical evidence. From the literature reviewed in the present study, there appeared to be little meta-analytic work examining the effects of only providing explicit English reading instruction for ESL students without literacy instruction in their native language. Such studies, when compared to dual instruction for example, could offer insights about the comparative benefits of each approach.

With the advent of the Reading First program, there is clearly a need for experimental studies with larger sample sizes where students are randomly assigned to Reading First and control groups. Unlike the present causal-comparative study, such experimental design could further limit the influence of variables in an effort to better highlight the effects or lack thereof, of Reading First strategies. It is interesting to note that following the present study, Congress mandated that the General Accounting Office conduct a study of the Reading First program (Education News, 2005). Reportedly, critics have alleged that Reading First failed to fulfill the intent of Congress and ignored other programs with demonstrated effectiveness.

A further area of need is longitudinal study of Reading First ESL participants. Due to the program's recency, there is understandably an absence of information on how these students fare in later school years. Issues that should be addressed would include the catch up rate of ESL students with their monolingual peers in reading achievement and whether their drop out rates and graduation rates are influenced or not by Reading First participation. Collier's (1992) discussion of methodological hazards in studying ESL students' achievement would appear particularly relevant to the consideration of studies of this nature.

Analyses were performed in this study to examine the predictability of reading measures on the Florida state achievement test (FCAT). Specifically, the GRADE and DIBELS were found to be significant predictors. However, only the spring 2005 administration of the DIBELS was used for analysis. Consistent with Reading First requirements, the DIBELS was administered four times during the school year. Efforts to determine when during the year the DIBELS is first able to predict subsequent FCAT scores would be beneficial. This would allow schools to begin intervention at the earliest possible date for those students most likely to fail the FCAT.

The present study used quantitative measures to examine explicit reading instruction for ESL students. Studies employing qualitative analyses could offer insight into the affective aspects of Reading First, both for students and teachers. Of particular interest would be the influence of Reading First strategies on the self image of students, their perceptions of success, and their motivation to persist in learning the literacy of a second language.

Due to the small sample sizes in this study and the limitations noted, generalization of findings would be tentative and limited to similar students in the school district where the data were collected. Further study of all ESL students who participate in Reading First schools would provide more reliable information about the benefits of the program.

As noted above, nearly half of the students participating in this study failed the 2005 FCAT reading test despite the progress they made between pre and posttest GRADE measures. This finding, together with the literature pertaining to second language acquisition, suggested that school districts should carefully examine the

efficacy of administering high stakes tests before grade 6 or later to ESL students who have only participated in immersion programming.

In the area of Collective Teacher Efficacy, this study suggested a need for more investigation into the effects of curricular approaches and programming such as Reading First where other variables are better controlled. For example, teacher experience and training if held constant might allow greater insight into the effects of curriculum and instruction. There is the obvious question that if levels of both or each independently increase does the degree of collective efficacy also rise? In addition, the concept of an *enabling* school structure as described by Hoy and Sweetland (2001) could also be relevant to the present findings. According to these authors, enabling schools essentially are characterized by procedures, rules, and a hierarchical authority system that provide teachers with a sense of power and autonomy. Their scale for measurement of enabling bureaucratic structures was found to be reliable. For future studies in the area of collective teacher efficacy, it would be beneficial to include a measure of enabling school structure in an effort to consider another possible and influential variable.

APPENDIX A

COLLECTIVE TEACHER EFFICACY SCALE-SHORT FORM

Directions: Please indicate the degree to which you agree or disagree with each statement below by circling the appropriate number to the right of each statement.

	Strongly disagree	Moderately disagree	Disagree slightly more than agree	Agree slightly more than disagree	Moderately agree	Strongly agree
	1	2	3	4	5	6
1. Teachers in this school are able to get through to difficult students.	1	2	3	4	5	6
2. Teachers here are confident they will be able to motivate their students.	1	2	3	4	5	6
3. Teachers in this school really believe every child can learn.	1	2	3	4	5	6
4. If a child doesn't want to learn, teachers here give up.	1	2	3	4	5	6
5. Teachers here don't have the skills needed to produce meaningful student learning.	1	2	3	4	5	6
6. These students come to school ready to learn.	1	2	3	4	5	6
7. Home life provides so many advantages the students here are bound to learn.	1	2	3	4	5	6
8. Students here just aren't motivated to learn.	1	2	3	4	5	6
9. The opportunities in this community help ensure these students will learn.	1	2	3	4	5	6

	Strongly disagree	Moderately disagree	Disagree slightly more than agree	Agree slightly more than disagree	Moderately agree	Strongly agree
	1	2	3	4	5	6
10. Learning is more difficult at this school because students are worried about their safety.	1	2	3	4	5	6
11. Drug and alcohol abuse in the community make learning difficult for students here.	1	2	3	4	5	6
12. Teachers in this school do not have the skills to deal with student disciplinary problems.	1	2	3	4	5	6

APPENDIX B  
TEACHER QUESTIONNAIRE

1. What is your gender?                      Male                      Female
2. How many years have you been teaching? \_\_\_\_\_
3. Have you completed *Reading First* training?      Yes                      No
4. What is the highest degree you have earned?  
    Bachelors  
    Master's  
    Specialists  
    Doctorate  
    Other \_\_\_\_\_
5. How many years have you been teaching at this school? \_\_\_\_\_
6. What grade do you teach? \_\_\_\_\_

APPENDIX C

PRINCIPAL PERMISSION FORM – READING FIRST SCHOOL

Date

Dear \_\_\_\_\_,

I am currently a graduate student at the University of Central Florida and am beginning research for a dissertation under the supervision of Drs. Larry Holt and Rick DuVall. The purpose of my study is to examine the benefits of *Reading First* strategies for Hispanic students. My intention is to compare pre- and post- GRADE test scores between two samples of Hispanic students. Those students in the *Reading First* school will serve as the experimental group and will be compared to a control group of non-*Reading First* students. Additionally, I plan to correlate DIBELS, FCAT, and GRADE scores to determine if these tests might have predictive validity for this population of students. As a secondary part of the study, the *Collective Teacher Efficacy Scale* (CTES) will be given to teachers in both *Reading First* and non-*Reading First* schools. Scores from the CTES will be correlated with the GRADE results.

Your school was selected due to its demographic characteristics, ESOL population, and participation in *Reading First*. Your permission is being requested to support the participation of a sample of students in grades two through five. To be selected for the study, the students must be Hispanic, have demonstrated a CALP level of 3 during the fall or an earlier administration of the BVAT, have both May 2004 and May 2005 GRADE scores available on Data Warehouse, not be identified as participants in Exceptional Student Education, show no previous retentions, and have had no past instruction in Spanish literacy. It is estimated that the sample size will be approximately 40 to 50 students. To participate, I would request that you release student identification numbers for your LEP population to allow selection of the sample and then obtain test scores from Data Warehouse. The only request I would make of your staff would be the release of DIBELS scores for the sample of students. For the second element of the study, your teachers would be mailed the 12-item *Collective Teacher Efficacy Scale* and a very brief questionnaire. This will require 10 minutes or less of their time and their consent would be requested for participation in the study. As the principal, you will not be asked to complete any forms or provide any additional information other than that described in this letter. Your staff would not be asked to participate in the sample selection.

The district's Research Oversight Committee has approved this research. Strict confidentiality will be maintained and all data pertaining to student scores and teacher responses will remain anonymous. Scores will be analyzed as group data. All data gathered for this study will be kept in a locked file with only the researcher having access.

I am excited about this research as findings could have positive educational implications for our students learning English as their second language. The results could also provide insight into the relationship between instructional methodology and teachers' beliefs about their school and teaching. If you choose to give permission or not to support your school's participation, please sign and return this letter to me. If you have

any questions, please feel free to call me at (239) 566-1675. Thank you very much for your consideration.

Sincerely,

L. Van Hylemon

I have read the description of the proposed study.

I give permission for the release of student identification numbers for the LEP population.

I request a copy of the data analysis.

I do not give permission for the release of any student identification numbers.

APPENDIX D

PRINCIPAL PERMISSION FORM – NONREADING FIRST SCHOOL

Date

Dear \_\_\_\_\_,

I am currently a graduate student at the University of Central Florida and am undertaking research for a dissertation under the supervision of Drs. Larry Holt and Rick DuVall. The purpose of my study is to examine the benefits of *Reading First* strategies for Hispanic students. My intention is to compare pre- and post- GRADE test scores between two samples of Hispanic students. Those students in the *Reading First* school will serve as the experimental group and will be compared to a control group of non-*Reading First* students. Additionally, I plan to correlate DIBELS, FCAT, and GRADE scores to determine if these tests may have predictive validity for this population of students. As a secondary part of the study, the *Collective Teacher Efficacy Scale* (CTES) will be given to teachers in both *Reading First* and non-*Reading First* schools. Scores from the CTES will be correlated with the GRADE results.

Your school was selected due to its demographic characteristics and ESOL population. Your permission is being requested to support the participation of a sample of students in grades two through five. To be selected for the study, the students must be Hispanic, have demonstrated a CALP level of 3 during the fall or earlier administration of the BVAT, have both May 2004 and May 2005 GRADE scores available on Data Warehouse, not be identified as participants in Exceptional Student Education, show no previous retentions, and have had no past instruction in Spanish literacy. It is estimated that the sample size will be approximately 40 to 50 students. To participate, I would request that you release student identification numbers for your LEP population to allow selection of the sample and then obtain test scores from Data Warehouse. For the second element of the study, your teachers would be mailed the 12-item *Collective Teacher Efficacy Scale* and a very brief questionnaire. This will require 10 minutes or less of their time and their consent would be requested for participation in the study. As the principal, you will not be asked to complete any forms or provide any additional information other than that noted in this letter.

The district's Research Oversight Committee has approved this research. Strict confidentiality will be maintained and all data pertaining to student scores and teacher responses will remain anonymous. All data gathered for this study will be kept in a locked file with only the researcher having access.

I am excited about this research as findings could have positive educational implications for our students learning English as their second language. The results could also provide insight into the relationship between curriculum and teachers' beliefs about their school and teaching. If you choose to give permission or not to support your school's participation, please sign and return this letter to me. If you have any questions, please feel free to call me at (239) 566-1675. Thank you very much for your consideration.

Sincerely,

L. Van Hylemon

I have read the description of the proposed study.

I give permission for the release of student identification numbers for the LEP population.

I request a copy of the data analysis.

I do not give permission for the release of any student identification numbers.

APPENDIX E  
TEACHER CONSENT FORM

Date

Dear Teacher:

I am a graduate student at the University of Central Florida and am currently working on a dissertation. As a part of my research, I am interested in the relationship between curriculum and collective teacher efficacy (CTE). CTE is defined as teachers' collective beliefs that they can positively influence student learning at their school. Such beliefs have been found to have a positive relationship with student achievement. A second part of my research will compare the reading achievement of Hispanic students in *Reading First* and non-*Reading First* schools. I am asking you to participate in the study by completing a brief questionnaire and the 12-item *Collective Teacher Efficacy Scale* (CTES). This will likely require less than 10 minutes of your time. You do not have to answer any question you do not wish to answer. Your responses will be kept confidential and your identity will be anonymous. I do not ask that you place your name on either the questionnaire or the CTES.

There are no anticipated risks, compensation, or other direct benefits to you as a participant in this study. You are free to withdraw your consent to participate in the study at any time.

If you have any questions about the study, please contact me at (239) 566-1675. My dissertation co-chairs are Dr. Larry Holt and Dr. Richard DuVall. Any questions or concerns you may have about research participants' rights may be directed to the UCFIRB office, University of Central Florida Office of Research, Orlando Tech Center, 12443 Research Parkway, Suite 207, Orlando, Florida 32826. The phone number is (407) 823-2901.

Please sign and return this letter in the enclosed envelope, along with the questionnaire and CTES if you agree to participate. A second copy of this letter is provided for your records. By signing this letter, you give me permission to compile your responses with those of other teachers and report the results anonymously in the final manuscript. Thank you very much for your consideration.

Sincerely,  
L. Van Hylemon

\_\_\_\_\_ I have read the description of the study and the procedures above.  
\_\_\_\_\_ I voluntarily agree to complete the questionnaire and CTES.  
\_\_\_\_\_ I would like to receive a copy of the final analysis of the data.  
\_\_\_\_\_ I would not like to receive a copy of the final analysis of the data.

\_\_\_\_\_/\_\_\_\_\_  
Participant / Date

APPENDIX F  
IRB APPROVAL LETTER



March 22, 2005

Larry Hylemon  
9596 Sussex Street  
Naples, FL 34109

Dear Mr. Hylemon:

With reference to your protocol #05-2490 entitled, "Collective Teacher and Reading Achievement for Hispanic Student in Reading First and Non-Reading First Schools in South West Florida" I am enclosing for your records the approved, expedited document of the UCFIRB Form you had submitted to our office. **The expiration date for this study will be 3/14/06.** Should there be a need to extend this study, a Continuing Review form must be submitted to the IRB Office for review by the Chairman or full IRB at least one month prior to the expiration date. This is the responsibility of the investigator. **Please notify the IRB when you have completed this study.**

Please be advised that this approval is given for one year. Should there be any addendums or administrative changes to the already approved protocol, they must also be submitted to the Board through use of the Addendum/Modification Request form. Changes should not be initiated until written IRB approval is received. Adverse events should be reported to the IRB as they occur.

Should you have any questions, please do not hesitate to call me at 407-823-2901.

Please accept our best wishes for the success of your endeavors.

Cordially,

*Barbara Ward*

Barbara Ward, CIM  
IRB Coordinator

Copy: IRB file

## LIST OF REFERENCES

- Assessment and Accountability Briefing Book, the Florida Comprehensive Assessment Test (2004). Retrieved October 28, 2004 from <http://www.firn.edu/doe/sas/fcat/pdf/fcataabb.pdf>.
- August, D. (2003). Supporting the development of English literacy in English language learners: Key issues and promising practices. Center for Research on the Education of Students Placed At Risk. Washington, DC: Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. ED474611)
- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research and Practice, 20*, 50-57.
- August, D., & Hakuta, K. (Eds.). (1997). *Improving schooling for language minority children: A research agenda*. Washington, D.C., National Academy Press.
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.
- Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Beaver, J. (1997). *Developmental Reading Assessment: Resource guide*. Glenview, IL: Celebration Press.
- Becker, W.C., & Gersten, R. (1982). A follow-up on Follow Through: The later effects of the Direct Instruction model on children in fifth and sixth grades. *American Educational Research Journal, 19*, 75-92.
- Brisbois, J.E. (1995). Connections between first and second-language reading. *Journal of Reading Behavior, 27*, 565-584.
- Brouwers, A., & Tomic, W. (2001). A test of the factorial validity of the Teacher Efficacy Scale. *Research in Education, 69*, 67-79.
- Calderon, M., Hertz-Larowitz, R., & Slavin, R.E. (1998). Effects of bilingual cooperative integrated reading and composition on students making the transition from Spanish to English reading. *Elementary School Journal, 99*. 153-165.

- Camilli, G., Vargas, S., Yurecko, M. (2003). Teaching children to read: The fragile link between science and federal education policy. *Education Policy and Analysis Archives, 11*. Retrieved June 24, 2005 from <http://epaa.asu.edu/epaa/vlln15/>.
- Carlo, M.S., August, D., McLaughlin, B., Snow, C.E., Dressler, C., Lippman, D. et al. (2004). Closing the gap: Addressing the vocabulary needs of English language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*, 188-216.
- Chomsky, N. (1959). Review of verbal behavior. *Language, 35*, 26-58.
- Collier, V. (1989). How long? A synthesis of research on academic achievement in a second language. *TESOL Quarterly, 3*, 509-528.
- Collier, V. (1992). A synthesis of studies examining long-term language minority student data on academic achievement. *Bilingual Research Journal, 16*, 187-212.
- Constantino, M. (1999). *Reading and second language learners: Research report*. Olympia, WA: Washington Office of the State Superintendent of Public Instruction. (ERIC Document Reproduction Service No. ED452706)
- Creswell, J. W. (2003). *Research design: qualitative, quantitative, and mixed methods approaches* (2<sup>nd</sup> ed). Thousand Oaks, CA: Sage Publications.
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.
- Cummins, J. (1979a). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some matters. *Working Papers on Bilingualism, 19*, 121-129.
- Cummins, J. (1979b). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*, 222-251.
- Cummins, J. (1999). BICS and CALP: Clarifying the distinction. U.S. Department of Education, Office of Educational Research and Improvement. Washington, DC. (ERIC Document Reproduction Service No. ED438551)
- Dale, R.L. (2005). The effects of teacher collaboration and school trust on collective efficacy. (Doctoral dissertation, Oklahoma State University). *Dissertation Abstracts International, 65*, 2474. (UMI No. 3115027)
- Denton, C.A. (2000). The efficacy of two English reading interventions in a bilingual education program. Unpublished doctoral dissertation, Texas A&M University.

- Denton, C.A., Anthony, J.L., Parker, R., & Hasbrouck, J.E. (2004). Effects of two tutoring programs on the English reading development of Spanish-English bilingual students. *The Elementary School Journal*, 104, 289-305.
- Dukes, S.S. (2000). The relationships among student academic and behavior characteristics, instructional strategies, teacher efficacy and student referral rate. *Dissertation Abstracts International* 60, (7-A), 2443. (UMI No. 9938320)
- Duran, E., & Carnine, D. (1999). *Leyendo en español: Components of beginning reading for first grade general education and at-risk limited English proficient students. Summary report of curriculum validation study conducted on Leamos español reading program.* Columbus, OH: SRA/McGraw-Hill.
- Duran, E., Shefelbine, J., Carnine, L., Maldonado-Colon, E., & Gunn, B. (2003). *Systematic instruction in reading for Spanish-speaking students.* Springfield, IL: Charles C. Thomas Publisher.
- Durgunoglu, A.Y. (2002). Cross-linguistic transfer in literacy development and implications for language learners. *Annals of Dyslexia*, 52, 189-204.
- Durgunoglu, A.Y., Nagy, W., & Hancia-Bhatt, G. (1993). Cross-language of phonological awareness. *Journal of Educational Psychology*, 85, 453-465.
- Education News.org (2005, October 4). Congress calls on GAO to examine Reading First. (Retrieved October 27, 2005 from <http://www.educationnews.org/congress-calls-on-gao-to-examine-reading-first.htm>)
- Ervin-Tripp, S.M. (1973). *Is second language learning like the first?* Paper presented at the annual meeting of Teachers of English to Speakers of Other Languages, Puerto Rico.
- Fillmore, L.W. (1985). When does teacher talk work as input? In S. Gass and C. Madden (Eds.), *Input in second language acquisition* (pp. 17-50). New York, NY: Newbury House Publishers.
- Fisher, K.B., & Cabello, B. (1981). *Predicting student success following transition from bilingual programs* (Report No. 161). Los Angeles: Center for the Study of Evaluation. (ERIC Document Reproduction Service No. ED222091)
- Fitzgerald, J. (1995). English-as-a-second language learners' cognitive reading processes: A review of research in the United States. *Review of Educational Research*, 65, 145-190.

- Florida Association of School Psychologists (FASP) (2003). *Position statement on Florida's third grade retention mandate*. Retrieved June 17, 2004, from <http://www.fasp.org>.
- Florida Department of Education (2000-2001). *English for Speakers of Other Languages. Language Enriched Pupil (LEP), Data Base and Program Handbook*. Retrieved August 30, 2005 from <http://www.firn.edu/doe>.
- Flynt, E.S., & Cooter, R.B. (1998). *Flynt-Cooter Reading Inventory for the Classroom, Third Edition*. Upper Saddle River, NJ: Prentice-Hall, Inc.
- Galloway, M.M. (2003). The impact of an intensive English reading intervention on pre-kindergarten, kindergarten and first grade Hispanic English language learners (Doctoral dissertation, Texas A&M University, 2003). *Dissertation Abstracts International*, 64, 4322. (UMI No. 3115027)
- Gerber, M., Jimenez, T., Leafstedt, J., Villaruz, J., Richards, C., & English, J. (2004). English reading effects of small-group intensive intervention in Spanish for K-1 English learners. *Learning Disabilities Research & Practice*, 19, 239-251.
- Gersten, R. (1985). Structured immersion for language minority students: Results of a longitudinal evaluation. *Educational Evaluation and Policy Analysis*, 7, 187-196.
- Geva, E. (2000). Issues in the assessment of reading disabilities in L2 children: Beliefs and research evidence. *Dyslexia*, 6, 13-28.
- Goddard, R. (2002). A theoretical and empirical analysis of the measurement of collective efficacy: The development of a short form. *Educational and Psychological Measurement*, 62, 97-110.
- Goddard, R., Hoy, W.K., & Hoy, A. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Educational Research Journal*, 37, 479-507.
- Goddard, R., Logerfo, L., & Hoy, W.K. (2004). High school accountability: The role of perceived collective efficacy [Abstract]. *Educational Policy*, 18, 403-425.
- Good, R.H., & Kaminiski, R.A. (Eds.). (2001). *Dynamic Indicators of Basic Early Literacy Skills* (5<sup>th</sup> ed.). Eugene, OR: Institute for the Development of Educational Achievement. Available: <http://dibels.uoregon.edu>.
- Gottardo, A. (2002). The relationship between language and reading skills in bilingual Spanish-English speakers. *Topics in Language Disorders*, 22, 46-70.

- Grabe, W. (1991). Current developments in second-language reading research. *TESOL Quarterly*, 25, 375-406.
- Gunn, B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *The Journal of Special Education*, 34, 90-103.
- Gunn, B., Smolkowski, K., Biglan, A., & Black, C. (2002). Supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school: A follow-up. *The Journal of Special Education*, 36, 69-79.
- Guskey, T.R. (1987). *Teacher efficacy, self-concept, and attitudes toward the implementation of mastery learning*. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC. (ERIC Document Reproduction Service No. ED281838)
- Hakuta, K., Butler, Y.G., & Witt, D. (1999). *How long does it take English language learners to attain proficiency?* Unpublished paper. Stanford, CA: Stanford University.
- Hudson, R.F., Smith, S.W., & Smith S.W. (2001). Effective reading instruction for struggling Spanish-speaking readers: A combination of two literatures. *Intervention in School & Clinic*, 37, 36-40.
- Just Read, Florida! (n/d.). Florida Department of Education. Retrieved June 22, 2005, from [http://www.justreadflorida.com/reading\\_first.asp](http://www.justreadflorida.com/reading_first.asp).
- Just Read, Florida! (2002). K-3 Reading Academy, *Presenter Guide*. Austin, TX: University of Texas, Texas Education Agency.
- Krashen, S. (2001). More smoke and mirror: A critique of the National Reading NRP report on fluency. *Phi Delta Kappan*, 83, 119-124.
- Krashen, S.D., & Terrell, T.D. (1983). *The natural approach: Language acquisition in the classroom*. Hayward, CA: Alemany Press.
- Laufer, B. (2001). *Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading?* Paper presented at the American Association of Applied Linguistics, St. Louis, MO.
- Leafstedt, J.M. (2002). Crossover of phonological processing skills: A study of Spanish-speaking students in two instructional settings. (Doctoral dissertation, University of California, Santa Barbara, 2003). *Dissertation Abstracts International*, 63, 3473. (UMI No. 3067321)

- Lenneberg, E. (1967). *Biological foundations of language*. New York, NY: John Wiley and Sons.
- Lopez, M.G., & Tashakkori, A. (2003). Utilizing two-way bilingual education for reducing the achievement lag of LEP students in primary grades: A longitudinal study. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED478290)
- Miller, K. (2003). English language learners and the No Child Left Behind Act. Mid-Continent Research for Education and Learning, Aurora, CO. (ERIC Document Reproduction Service No. ED482752)
- Muñoz-Sandoval, A.F., Cummins, J., Alvadero, C.G., & Rueda, M.C. (1998). *Bilingual Verbal Ability Tests*. Itasca, IL: Riverside Publishing Company.
- National Center for Educational Statistics (2002). Public school students, staff, and graduate counts by state, school year 2000-2001. Washington, DC: U.S. Department of Education.
- National Institute of Child Health and Human Development (2000a). Report of the National Reading NRP. Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Retrieved January 27, 2005 from <http://www.nichd.nih.gov/publications/nrp/smallbook.htm>.
- National Institute of Child Health and Human Development. (2000b). Report of the National Reading NRP. Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implication for reading instruction: Reports of the subgroups. Retrieved January 27, 2005 from <http://www.nichd.nih.gov/publications/nrp/smallbook.htm>.
- Nelson, M.L. (2003). Assessing the early literacy skills of young English learners: Use of DIBELS in Spanish. *Dissertation Abstracts International*, 64, 2825. (UMI No. 3102181)
- Nye, B., Konstantopoulos, S. & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26, 237-257.
- Perez, E. (1981). Oral language competence improves reading skills of Mexican-American third graders. *The Reading Teacher*, October, 24-27.

- Pinnell, G.S., Pikulski, J.J., Wixson, K.K., Campbell, J.R., Gough, P.B., & Beatty, A.S. (1995). *Listening to children read aloud*. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education. (ERIC Document Reproduction Service No. ED378550)
- Ramirez, J.D. (1992). Executive summary of the final report: Longitudinal study of structured English immersion strategy, early-exit and late-exit transitional bilingual education program for language-minority children. *Bilingual Research Journal: The Journal of the National Association for Bilingual Education*, 16, 1-62.
- Ross, J.A., Hogaboam-Gray, A., & Gray, P. (2003). *The contribution of prior student achievement and school processes to collective teacher efficacy in elementary schools*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED479719)
- Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80, 1-28.
- Shames, R. (1998). The effects of a community language learning/comprehension processing strategies model on second language reading comprehension. Unpublished doctoral dissertation, Florida Atlantic University.
- Slavin, R.E., & Cheung, A. (2003). Effective reading programs for English language learners. A best-evidence synthesis. Center for Research on the Education of Students Placed At Risk, Washington, DC: Institute of Education Sciences. (ERIC Document Reproduction Service No. ED483006).
- Slavin, R.E., & Cheung, A. (2004). How do English language learners learn to read? *Educational Leadership*, 61, 52-58.
- Slavin, R.E., & Madden, N. (1999). Effects of bilingual and English as a second language adaptations of Success for All on the reading achievement of students acquiring English as a second language. *Journal of Educators for Students Placed At Risk*, 4, 393-416.
- Snow, C.E., Burns, M.S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- SPSS, Inc. (2002). *SPSS graduate pack 11.5 for Windows* (11.5 ed.). Chicago: SPSS, Inc.
- Starko, A.J., & Schanck, G.D. (1989). Perceived need, teacher efficacy, and teaching strategies for the gifted and talented. *Gifted Child Quarterly*, 33, 118-122.

- Stuart, M. (1999). Getting ready for reading. Early phoneme awareness and phonics training improves reading and spelling in inner-city second language learners. *British Journal of Educational Psychology*, 69, 587-605.
- Torgesen, J.K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, 40, 7-26.
- Torgesen, J.K., Wagner, R.K., Rashotte, C.A., Rose, E., Lindamood, P., Lindamood, P., et al. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology*, 91, 579-593.
- Thomas, W. P., & Collier, V.P. (1997). *School effectiveness for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education.
- Thomas, W. P., & Collier, V. P. (2002). A national study of school effectiveness for language minority students' long-term achievement. Washington, DC: Office of Educational Research and Improvement. Retrieved July 2, 2004, from [http://www.crede.ucsc.edu/research/llaa/1.1\\_final.html](http://www.crede.ucsc.edu/research/llaa/1.1_final.html).
- Tschannen-Moran, M., Hoy, A., & Hoy, W.K. (1998). Teacher efficacy: its meaning and measure. *Review of Educational Research*, 68, 202-248.
- U.S. Department of Education. (1999). *The NAEP 1998 report card for the nation and the states*. Office of Educational Research and Improvement. National Center for Education Statistics. Washington, DC. Retrieved June 26, 2004 from <http://nces.ed.gov/pubsearch/pubsinfo>.
- U. S. Department of Education (2004). *No Child Left Behind*. Retrieved June 26, 2004 from <http://www.ed.gov/nclb/landing.jhtml>.
- Van Wagenen, M.A., Williams, R.L., & McLaughlin, T.F. (1994). Use of assisted reading to improve reading rate, word accuracy, and comprehension with ESL Spanish-speaking students. *Perceptual and Motor Skills*, 79, 227-230.
- Wheatley, K.F. (1998). The relationship between efficacy beliefs and reform-oriented Mathematics teaching: three case studies. *Dissertation Abstracts International*, 58, (9-A), 3451. (UMI No.9938320)
- Williams, K.T. (2002). *Group reading assessment and diagnostic evaluation (GRADE)*. Circle Pine, MN: AGS Publishing.