

Florida State University Libraries

Electronic Theses, Treatises and Dissertations

The Graduate School

2005

The Relationship Between the Implementation Practices of School Principals and Student Achievement in Reading

Stephen M. Nettles



THE FLORIDA STATE UNIVERSITY

COLLEGE OF EDUCATION

THE RELATIONSHIP BETWEEN THE
IMPLEMENTATION PRACTICES OF SCHOOL PRINCIPALS
AND STUDENT ACHIEVEMENT IN READING

By

STEPHEN M. NETTLES

A Dissertation submitted to the
Department of Educational Leadership and Policy Studies
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded
Fall Semester, 2005

The members of the Committee approve the dissertation of Stephen M. Nettles defended on September 12, 2005.

Joseph Beckham
Professor Directing Dissertation

Stephanie al Otaiba
Outside Committee Member

Carolyn Herrington
Committee Member

Michael Biance
Committee Member

Laura Hassler
Committee Member

Approved:

Joseph Beckham, Chairperson
Department of Educational Leadership and Policy Studies

The Office of Graduate Studies has verified and approved the above named committee members.

This is for my beautiful family. To my wife Char, there is no way that this would have happened without your love and support. Everything I am, or will ever be, I owe to you. Thank you, baby. To Matthew, I feel you with me everyday. You are forever my promise of better things to come. To MadelineMarie, you repaired my heart with a touch of your hand. Your fiery strength impresses me everyday. I cannot wait to see the woman you become. And, to Stone, you are the completing piece of my life, without which I would never be right. You are the perfect combination of lamb and lion; a gentle soul with the heart of a warrior. Maintain that balance and you will be a great man. I love you all.

ACKNOWLEDGMENTS

Throughout this process, there are people who provided support and assistance and for which I extend my deepest thanks. To Dr. Joe Beckham, my sincere appreciation for your time and talent and for helping me maintain my focus. I will always be thankful for the guidance of my doctoral committee — Dr. Carolyn Herrington, Dr. Laura Hassler, Dr. Stephanie Al-Otaiba, and Dr. Mike Bianche — for their knowledge and dedication. In particular, I want to thank Dr. Mike Bianche for his friendship and for serving as my mentor throughout this process. As well, my deepest thanks go to my friends Yaacov Petscher and Dr. Karen Nicholas for their invaluable assistance on this project. TMV!

TABLE OF CONTENTS

List of Tables.....	v
List of Figures.....	vii
Abstract.....	viii
1. INTRODUCTION.....	1
Statement of the Problem.....	3
Purpose of this Study.....	4
Research Questions and Hypotheses.....	5
Conceptual Framework.....	7
Exogenous Variables of Student Achievement	
Endogenous Variables of Student Achievement	
Principal as Instructional Leader	
Principal Behaviors and Student Achievement	
Methods and Procedures.....	11
Selected Data	
Statistical Analysis	
Definition of Terms.....	13
Effective Reading Programs	
Oral Reading Fluency	
Limitations to the Study.....	14
Delimitations to the Study.....	14
2. REVIEW OF THE LITERATURE.....	15
Exogenous Variables of Student Achievement.....	15
Endogenous Variables of Student Achievement.....	18
Education Production Function Research	
Effective Schools Research	
Instructional Leadership	
Clear and Focused Mission	
Safe and Orderly Environment	
Climate of High Expectations	
Frequent Monitoring of Student Progress	
Positive Home-School Relations	
Opportunity to Learn and Student Time on Task	
Additional Endogenous Factors of Student Achievement	
School/Class Size	

Instructional Quality	
Teacher Qualifications	
School Climate	
The Principal as Instructional Leader.....	27
Safe and Orderly Environment	
Mission and Vision	
Stakeholder Involvement	
Monitoring School Progress	
Instructional Focus	
High Expectations for Student Performance	
Professional Development	
Principal Behaviors and Student Achievement.....	31
3. METHODS.....	35
Purpose and Research Questions.....	35
Research Methods.....	40
Sampling Procedures	
Principal Population	
Student Population	
Data Selection	
Measures and Reliability.....	42
Core Reading Program	
Professional Development	
Leadership/Organizational Practices (Leadership)	
Assessment	
Intervention	
Validation Process	
Data Summary and Analysis Procedures.....	52
4. FINDINGS.....	56
Research Data.....	56
Principal Implementation Questionnaire (PIQ) Responses	
Population Subgroup ORF Data	57
Socio-economic Status (SES)	
Gender	
Ethnicity	
Limited English Proficiency (LEP)	
Students Qualifying for Exceptional Student Education (ESE)	
Three-Level Growth Curve Model.....	65
Model Formulation	

Analysis of Variance (ANOVA) Model	
Level 1 Model	
Level 2 Model	
Level 3 Model	
Research Questions and Hypotheses.....	73
Research Question #1	
Research Question #2	
Research Question #3	
Research Question #4	
Research Question #5	
Research Question #6	
 5. DISCUSSION AND CONCLUSIONS.....	 84
3-Level Growth Curve Model.....	84
Principal Effects on ORF Growth.....	84
Principal Effects on ORF Growth in Relation to the Total	
Population.....	85
Principal Effects on ORF Growth in Relation to SES	85
Principal Effects on ORF Growth in Relation to Gender.....	86
Principal Effects on ORF Growth in Relation to Students	
with Limited English Proficiency.....	89
Principal Effects on ORF Growth in Relation to	
Students Qualifying for Exceptional Student Education.....	91
Effect Size.....	92
Conclusions and Recommendations for Future Research.....	93
Self-Report Limitations.....	94
Direct Effects of Principal Implementation Practices.....	95
Implications for Instructional Leadership.....	96
 APPENDIX.....	 98
 REFERENCES.....	 111
 BIOGRAPHICAL SKETCH.....	 127

LIST OF TABLES

1. Descriptive Statistics for PIQ Results.....	46
2. Summary of Model Fit Indices for CFA Iterations.....	48
3. Estimated Structure Coefficients for PIQ Constructs.....	50
4. Correlations for PIQ Dimensions.....	52
5. Descriptive Statistics for Retained PIQ Items by Construct.....	58
6. Means and Standard Deviations of Socio-economic Status as Indicated by Free/Reduced Lunch Eligibility.....	60
7. Means and Standard Deviations by Gender.....	61
8. Means and Standard Deviations by Ethnicity.....	62
9. Means and Standard Deviations by LEP Classification.....	63
10. Means and Standard Deviations by ESE Classification.....	64
11. Final Estimation of Variance Components for ANOVA Model.....	67
12. Final Estimation of Fixed Effects for Level 1.....	68
13. Final Estimation of Level 1 and Level 2 Variance Components.....	68
14. Final Estimation of Fixed Effects for Level 2 (Initial Intercept).....	70
15. Final Estimation of Fixed Effects for Level 2 (Slope).....	71
16. Final Estimation of Fixed Effects for Level 1 and Level 2 Variance Components.....	72
17. Final Estimation of Fixed Effects for Level 3 (Intercept).....	75
18. Final Estimation of Fixed Effects for Level 3 (Slope).....	77

19. HLM Interactions by PIQ Domain and Student Group.....	82
20. Final Estimation of Fixed Effects for Level 1 and Level 2 Variance Components.....	83

LIST OF FIGURES

1. Analysis Diagram for Three-level Growth Curve Model.....	55
2. Slope of ORF Growth for All Student Subgroups.....	74
3. Slope of ORF Growth Based On Intervention Responses for Total Population.....	86
4. Slope of ORF Growth Based On Leadership Responses for SES Subgroup.....	87
5. Slope of ORF Growth Based On Assessment Responses for SES Subgroup.....	88
6. Slope of ORF Growth Based On Leadership Responses for Gender Subgroup.....	89
7. Slope of ORF Growth Based On Professional Development Responses for Gender Subgroup.....	90
8. Slope of ORF Growth Based On Assessment Responses for Gender Subgroup.....	91
9. Slope of ORF Growth Based On Leadership Responses for LEP Subgroup.....	92
10. Slope of ORF Growth Based On Professional Development Responses for LEP subgroup.....	93
11. Slope of ORF Growth Based On Leadership Responses for ESE Subgroup.....	94

ABSTRACT

This exploratory study investigated the effects of principal implementation behaviors, specifically those influencing early reading programs, on student growth in oral reading fluency (ORF) of first grade students in Florida's *Reading First* schools. This study examined the relationship between the principal's role in the implementation of effective reading programs and student achievement in reading. Data consisted of (a) responses to the Principal Implementation Questionnaire (PIQ), and (b) student reading achievement records as measured by DIBELS Oral Reading Fluency (ORF) scores. A three-level growth curve model was used to determine the amount of student-level variance that can be explained by the five dimensions measured by the PIQ. Positive, statistically significant relationships were found between selected principal implementation practices (as measured by the PIQ) and student growth in oral reading fluency. This study concluded that certain principal behaviors associated with implementing effective reading programs display a direct, linear relationship to student achievement. Further, while these relationships account for a small proportion of the total student achievement variability in reading, they are of sufficient magnitude to be of interest and additional investigation. Findings should be used to develop professional development opportunities tailored to teach effective implementation of reading interventions for struggling readers across all subgroups to principals. Future research also has the potential to discriminate between students with differing classifications of disabilities in relation to principal effects on reading.

CHAPTER 1

INTRODUCTION

There is little argument regarding the difficulty of a school principal's job. It is apparent to anyone that has spent time in an American public school that the school principal wears many hats during the typical day. Presently, principals are expected to be business and personnel managers, public relations experts and social workers, in addition to their responsibility for effectively overseeing and driving effective instructional programs. While the processes associated with teaching and learning are the foundation of a school's core mission, the result of the complex job descriptions of principals is that managerial and administrative duties often leave little time for effective instructional leadership (Cuban, 1988; Deal & Peterson, 1990; Graham, 1997; Stronge, 1988).

Leading up to the 1970s, it was commonly understood that principals dealt primarily with building-level management and student discipline. It was in this period that their roles began evolving because of emerging research on effective schools (Brookover & Lezotte, 1982; Deal & Peterson, 1990; Edmonds, 1979). This body of work showed that the responsibilities and actions of principals were linked directly to student achievement (Brookover & Lezotte, 1982; Edmonds, 1979; Weber, 1971). The concept of instructional leadership is a relatively new understanding that emerged in the early 1980s, calling for principals' primary responsibilities to be shifted from managerial and administrative to those focused on instructional matters. This shift was influenced largely by researchers who found that effective schools usually had principals who practiced effective instructional leadership (Brookover & Lezotte, 1982). In the 1990s, "attention to instructional leadership seemed to waver, displaced by discussions of school-based management and facilitative leadership" (Lashway, 2002, p.1).

While most researchers and practitioners would agree that instructional leadership is essential to the development of effective schools, it is seldom practiced with consistency. For example, principals have been shown to devote only one-tenth of their total time on the job to instructional issues (Stronge, 1988). Today, principals are challenged to balance their role as manager/administrator and instructional leader.

Unfortunately, as Herriot and Firestone (1984) found, managerial tasks associated with the principalship are often unrelated to those that promote effective and lasting school change. Less emphasis is given to instructional leadership due to (a) the lack of comprehensive training as instructional leaders, (b) lack of time to initiate or complete instructional activities, (c) increased paper work and (d) the community's expectation that the principal's role is primarily that of a manager (Elmore, 2000; Fullan, 1991, 1997; Graham, 1997). New standards of accountability for student performance have pushed principals into bearing primary responsibility for student achievement for the first time in history. Thus, this role of instructional leadership to specifically target and increase student learning is new to most principals (Lambert, 1998).

Moreover, there is a difference between what job responsibilities principals value and those they actually perform (Graham, 1997). In a study conducted by Graham (1997) involving approximately 500 school principals, more than two-thirds considered themselves administrators or managers, while only one-fourth considered themselves instructional leaders. The average principal in this study reported spending a majority of the work day on administrative duties. Further, principals reported spending less than five hours per week on matters associated with instructional leadership.

The current situation is further complicated because schools throughout our country are experiencing unprecedented challenges that can be attributed to many forces, both internal and external. American schools have never been asked to accomplish as much, both academically and socially, as expected today. More and more, the focus in contemporary education is on the individual student. Thus, a vast majority of the consequences of education accountability rests on the individual school. Nationwide, schools are faced with the challenges associated with limited/disparate funding, changes in accountability standards and low student achievement. As Elmore (2000) stated:

Schools are being asked by elected officials-policy leaders to do things they are largely unequipped to do. School leaders are being asked to assume responsibilities they are largely unequipped to assume and the risk and consequences of failure are high for everyone, but especially high for children. (p.2).

Statement of the Problem

One of the primary indicators of overall student achievement is the ability to read and comprehend text. As such, reading is the focus of many national, state and local education accountability initiatives. Educators, parents and national leaders recognize that the limited reading skills of many children across our nation are an issue of national importance. The 2003 National Assessment of Educational Progress (NAEP) indicated that approximately 40% of fourth grade students cannot read at a basic level (NCES, 2003). When considering subgroups of students, almost 70% of low-income fourth grade students and approximately 50% of students living in urban environments cannot read at basic levels. In addition, student performance trends over a 10-year period show insignificant progress for average readers and even less progress for the lowest-performing readers (U.S. Department of Education, 2002).

There are many aspects of principal leadership that apply to the effort to improve student reading achievement. Researchers have investigated relationships between the behavior of school principals and safe and orderly school environments (Sammons, Thomas & Mortimore, 1997; Scheurich, 1998), school goals/vision that focus on high levels of student learning (Cotton, 2000; Johnson & Asera, 1999; Renchler, 1992), high expectations for student learning (Butler, 1997; Scheurich, 1998), positive and supportive school climate (Hallinger, Bickman & Davis, 1996), parent and community outreach (Bartell, 1990) and staff empowerment (Cotton, 1992; Gullatt & Lofton, 1996). More specifically, research has focused on leadership concerns such as instructional issues (Bamburg & Andrews, 1991), classroom observation and feedback to teachers (Heck, 1992; Leithwood & Jantzi, 2000; Spillane, Halverson & Diamond, 2001), professional development opportunities (Wagstaff, Melton, Lasless & Combs, 1998), protecting instructional time (Evans & Teddlie, 1995; Graham, 1997) and monitoring student progress by staff (Leithwood & Jantzi, 2000; Leitner, 1994).

Hallinger et al. (1996) found that principal leadership itself is influenced by both personal and contextual variables (socio-economic status [SES], parental involvement and gender). More specifically, these researchers explored the effects of school principals on the reading achievement of elementary school students. While finding no direct effects

of principal behaviors on reading outcomes, this study did identify some significant indirect effects of principal practice on student achievement through actions that shape school learning climate (Hallinger et al.). However, despite these and other investigations that illustrate the indirect impact of effective school leadership, there has been little inquiry that has focused directly on relationships between the instructional leadership provided specifically by elementary principals and student performance levels in reading.

Purpose of this Study

In 1998, the U.S. Department of Education in collaboration with the U.S. Department of Health and Human Services charged the National Academy of Sciences to investigate strategies for the prevention of reading difficulties and to determine the efficacy of reading interventions for struggling readers in the early grades. The body of literature on the topic was reviewed by a committee of reading experts. Among other findings, the committee determined that effective instruction in the early grades requires highly-qualified teachers that receive ongoing training and support in reading. The committee also found that effective instructional organizations must take into account the alignment of facilities, materials and services, as well as school environment, when determining appropriate instructional practices in reading (Snow, Burns & Griggin, 1998).

Reading First is a federal program authorized under Title I, Part B, Subpart 1 of the *No Child Left Behind Act of 2001* (NCLB, 2002) as a national initiative to improve the reading skills of students in kindergarten through third grade. *Reading First* is focused on effective classroom instruction supported by scientifically-based reading research. The program provides a large amount of funding (approximately \$1 billion for FY 2005) to states for activities targeting reading improvement in young students (USDOE, 2005). *Reading First* offers the opportunity for every state to receive funds, 20% of which can be used at the state level, while the remainder must be distributed to local education agencies (LEAs).

The implementation of *Reading First* requirements places additional responsibilities on school leadership, especially principals. *Reading First*, like NCLB, carries with it high-stakes consequences for failure in return for the funding it provides to

schools. Schools must set annual reading achievement goals that will facilitate growth to meet *Reading First* requirements of all students reading on grade level by 2013.

In Florida, schools that did not meet their progress goals at the end of the first year (2003-2004) were identified for intervention. Personnel from the state technical assistance team met with district representatives to formulate a technical assistance and support plan for any school that did not meet their achievement goals after the first year. The same procedure was followed at the end of the second year of implementation (2004-2005). If a school does not meet its achievement goals for reading at or above grade level by the end of the third year of *Reading First* implementation (2005-2006), the school district will then be asked to make a determination about whether the school should be continued as part of their *Reading First* grant for the next three years. Therefore, it is critical that each *Reading First* principal makes implementing effective reading programs a high priority, or risk losing significant funding (a minimum of \$40,000 per school, per year; FDOE, 2002).

The purpose of this study was to determine the relationship between the principal's role in the implementation of effective reading programs and student achievement in reading. In this study, effective reading programs were defined as those approved by the State of Florida for use in schools receiving *Reading First* program funding. Although research on instructional leadership and student achievement is prevalent, knowledge about the effects of principal leadership and its effects on reading achievement has not been comprehensively investigated. Specifically, this study has established a more focused understanding of the specific role that principals' implementation behaviors play in improving elementary student achievement in reading. It also addressed the disparity between the current practices of principals and the practices that they should be implementing to address student achievement in reading by analyzing principals' self-responses to a reading program implementation questionnaire.

Research Questions and Hypotheses

The following research questions and relative hypotheses were addressed in this study:

1. Is there a significant relationship between defined elements of principals'

implementation of effective reading programs and growth in student reading achievement in the total population of elementary students identified during the 2004-2005 academic year?

Null Hypothesis (H_{01}): There is no statistically significant difference in the growth trajectories of student reading achievement of students in schools where the principal displays high levels of implementation of effective reading programs versus those displaying lower levels.

Research Hypothesis (H_{a1}): There is a statistically significant difference in the growth trajectories of student reading achievement of students in schools where the principal displays high levels of implementation of effective reading programs versus those displaying lower levels.

2. Is there a significant difference between the reading achievement of first grade students of low socio-economic status and students of higher socio-economic status in relation to principal effects?

Null Hypothesis (H_{02}): There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to principal effects.

Research Hypothesis (H_{a2}): There is a statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to principal effects.

3. Is there a significant difference between the reading achievement of first grade students of different gender in relation to principal effects?

Null Hypothesis (H_{03}): There is no statistically significant difference in the growth trajectories of male and female students in relation to principal effects.

Research Hypothesis (H_{a3}): There is a statistically significant difference in the growth trajectories of male and female students in relation to principal effects.

4. Is there a significant difference between the reading achievement of first grade ethnic minority and non-minority students in relation to principal effects?

Null Hypothesis (H_{04}): There is no statistically significant difference in the growth trajectories of first grade ethnic minority and non-minority students in relation to principal effects.

Research Hypothesis (Ha₄): There is a statistically significant difference in the growth trajectories of first grade ethnic minority and non-minority students in relation to principal effects.

5. Is there a significant difference between the reading achievement of first grade students with limited English proficiency and students that are fluent in English in relation to principal effects?

Null Hypothesis (Ho₅): There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to principal effects.

Research Hypothesis (Ha₅): There is a statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to principal effects.

6. Is there a significant difference between the reading achievement of first grade students that qualify for exceptional student education and students that have no disability classification in relation to principal effects?

Null Hypothesis (Ho₆): There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to principal effects.

Research Hypothesis (Ha₆): There is a statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to principal effects.

Conceptual Framework

Exogenous Variables of Student Achievement

Throughout the body of literature on educational leadership, many variables have been identified as contributing to student achievement. The typical student achievement research predating the 1990s examined the impact of exogenous (out-of-school) variables on student achievement. Early studies in this area often focused on the background environment of individual students, citing the importance of these demographic variables on overall student achievement. For example, studies frequently suggested that students displaying higher socio-economic status (SES) would perform better in school than those

from less affluent backgrounds. Similarly, student-level variables such as innate intellectual capacity and motivation to learn are often identified as having strong effects on student performance (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld & York, 1966; Jencks, 1972, 1977; van de Grift & Houtveen, 1999). Conversely, endogenous (in school) variables are often cited as weak predictors of student achievement.

Endogenous Variables of Student Achievement

In his paper *Investment in Human Capital*, Schultz (1961) posited that investing in the educational attainment of workers (human capital) would produce a more robust U.S. economy than investing in physical capital, such as factories and machinery. Becker (1964) went on to suggest that human capital could be accumulated through investments in education and training opportunities. As educational institutions are the main resource for providing these types of opportunities, school researchers began to look for school-level factors that could be improved. This line of inquiry produced a body of education production function research which focused primarily on school spending and the resources that those funds can purchase.

In a meta-analysis of the body of research literature associated with educational spending, Hanushek (1989, 1991, 1997) summarized studies that had been produced since the release of the *Coleman Report* (1966). He analyzed 38 studies conducted from the 1960s through the 1980s to investigate the possible relationships between educational inputs and student outcomes. A total of 187 production function analyses were identified measuring seven variables, including (a) expenditures per pupil, (b) the experience of teachers, (c) the educational attainment of teachers, (d) the salaries paid to teachers, (e) the ratio of students to teacher, (f) factors related to school leadership and (g) facility-related factors.

Hanushek (1989) found that there was a very small percentage of positive and statistically significant correlations between economic inputs and student achievement outcomes for each of the seven variables. He concluded by stating that there was “no strong or systematic relationship between school expenditures and student performance” (p. 47). In other words, he reported that school characteristics and resources had an insignificant impact on student performance.

Effective schools research has also added to the understanding of the endogenous variables of student achievement. This research differed from Hanushek's work because the identified variables associated with student achievement could not necessarily be purchased. This area of research focused more on school-level environmental factors and less on traditional education resources.

The results of the original effective schools research and those of later studies on the topic have identified elements that comprise a successful school (Edmonds, 1979; Brookover & Lezotte, 1982; Lezotte, 2001). Effective schools research gained prominence in the 1980s with the release of the correlates of effective schools. The correlates include (a) instructional leadership, (b) clear and focused mission, (c) safe and orderly environment, (d) climate of high expectations, (e) frequent monitoring of student progress, (f) positive home-school relations, and (g) opportunity to learn and student time on task (Lezzote, 2001). These elements are now widely recognized as necessary for school success.

The Principal as Instructional Leader

Successful attainment of learning goals is associated with effective schools (Brookover & Lezotte, 1982; Edmonds, 1979; Marshall & Greenfield, 1985). There is a foundation of research on effective schools that suggests principals have considerable impact upon student's achievement through effective monitoring of the instructional process (Brookover & Lezotte, 1982; Edmonds, 1979; Lezzote, 2001). Because school principals are keys to successful schooling, understanding the ways in which they deal with existing problems in their schools and their ability to address these problems in light of current educational reforms becomes imperative. Even more, evidence suggests that principal leadership, as mediated through the development of these previously listed school level conditions and processes, has an effect on student learning (Hallinger & Heck, 1996).

Research has also been consistent in showing that school principals' leadership behavior is an essential factor for their effectiveness, as well as the effectiveness of their school (Cheng, 1994; Hallinger & Heck, 1999; Hoy & Miskel, 2001). A school's capacity for change stems from a combination of factors that enable administrators to problem-solve and implement decisions that help students learn. Aligned with this view,

various research studies have investigated the perception of the principal as a strong leader and the association with effective articulation of the school's mission, the creation of a safe learning environment and the implementation of appropriate instructional improvements (Edmonds, 1979; Berends, 2002; Clark & Clark, 2001).

Principal Behaviors and Student Achievement

Despite the large body of literature on educational leadership, the causal relationship between principal behaviors and student achievement remains unclear at best (Hallinger et al., 1996; Witziers, Bosker & Kruger, 2003). Furthermore, much of the research conducted on this potential relationship consists of over-simplified methodologies relying on bivariate statistical models (Hallinger & Leithwood, 1994; Hallinger et al., 1996; Witziers et al., 2003). This type of research generally seeks to establish a direct link between the actions of principals and the achievement of their students. Scholars have, for the most part, discarded the view that principals have a direct effect on student learning and have, instead, adopted the notion that principals have an impact on student achievement through their interactions with teachers and their efforts in shaping and managing the educational environment (Hallinger & Leithwood, 1994; Heck, 1993; Hallinger et al., 1996; Witziers et al., 2003). However, in a meta-analysis of 70 recent classroom and leadership studies, Waters, Marzano and McNulty (2003) investigated whether the quality of leadership has a significant direct relationship to student achievement and what specific leadership responsibilities and practices have the greatest impact. They concluded that school leadership is an important variable, as it correlates positively with student achievement.

Hallinger et al. (1996) suggest that previous research in this area has been weakened by the use of oversimplified methodological approaches; therefore, future research designs will be strengthened if they include (a) sufficient sample sizes, (b) theoretically defensible models, (c) reliable data collection instruments and (d) sophisticated data analysis tools. The use of these elements should enable researchers to better capture the indirect effects of principal behaviors on student achievement (Hallinger et al., 1996; Witziers et al., 2003). In an attempt to measure the potential affects of Florida *Reading First* principals' implementation behaviors on elementary

student reading achievement, this present study will incorporate all of these suggested elements.

Methods and Procedures

This study examined the relationship between the principal's role in the implementation of effective reading programs and student achievement in reading. Data used in this study to address the research questions and hypotheses consisted of (a) responses to the Principal Implementation Questionnaire (PIQ) and (b) student reading achievement records.

Selected Data

The PIQ is a 105-item Web-based inventory that was given to Florida school principals to inquire about the implementation of *Reading First* activities in their school during the 2004-2005 academic year. There are five dimensions of the PIQ that were self-reported in approximately 25 minutes by school principals: (a) Core Reading Program, (b) Professional Development, (c) Leadership/Organizational Practices, (d) Assessment and (e) Intervention. These five dimensions were selected and developed based on priorities set forth by state and federal *Reading First* policies, as well as a research-based understanding of the school-level elements needed to promote student achievement in reading. Further, each of these dimensions is focused on elements of overall reading program implementation that can be influenced by principal behaviors. Internal consistency analyses and confirmatory factor analyses were conducted to determine reliability and validity of the individual items within the PIQ. Chapter Three will include a detailed discussion of this process.

The student achievement data for this study included reading achievement as measured by the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessment (Good & Kaminski, 2002). DIBELS scores were analyzed for students in first grade attending Florida *Reading First* schools in the 2004-2005 academic year. DIBELS yields five sub-categories that are analyzed according to the grade levels in which they are administered: (a) Initial Sound Fluency, kindergarten only; (b) Letter Naming Fluency, kindergarten and first grade; (c) Phoneme Segmentation Fluency, kindergarten and first grade; (d) Nonsense Word Fluency, kindergarten, first and second grade; and (e) Oral

Reading Fluency, first, second and third grade.

For the purposes of this study, only DIBELS Oral Reading Fluency (ORF) data were used to measure the reading achievement of first grade students in *Reading First* schools. The rationale supporting this choice stems from research that consistently identifies student scores on assessments of oral reading fluency as strong predictors of overall reading comprehension, which is the benchmark for reading acquisition and the end goal for reading instruction (National Reading Panel, 2000; Snow, Burns & Griffin, 1998). Another compelling reason to use ORF scores is the fact that this measure is considered as one of the more stable instruments across a variety of assessment situations.

Statistical Analysis

Data were summarized and reported according to (a) the scores on the five sub-scale dimensions of the PIQ and (b) the four instances of the ORF subtest collected by the Florida Center for Reading Research (FCRR) from all *Reading First* schools in Florida during the 2004-2005 academic year. In a partnership with the Florida Department of Education, FCRR provides technical assistance and support to all districts and schools receiving a *Reading First* award. In addition, FCRR conducts basic research on reading, reading growth, reading assessment and reading instruction that will contribute to the scientific knowledge of reading.

A three-level growth curve model was used to determine the amount of student-level variance that can be explained by the five dimensions measured by the PIQ. Level 1 of this model tested the growth of ORF scores over time for each of the four assessments at the individual level. Level 2 modeled selected student-level variables, including gender, socio-economic status (SES)—as determined by free and reduced lunch eligibility, ethnicity, whether students have limited English proficiency and disability status. Level 3 modeled variables associated with principal. Specifically, principals' responses on the five dimensions of the PIQ were used to define and categorize principal behavior in relation to the implementation of effective reading programs.

Definition of Terms

Effective Reading Programs

The concept of effective reading programs is aligned and consistent with the scientific knowledge base in reading that underlies the *Reading First* component of the *No Child Left Behind Act of 2001*. For the purposes of this study, an effective reading program was defined as one that met the requirements of the Florida Department of Education (FDOE) requirements for use in *Reading First* schools, which state that: (a) *Reading First* schools must increase the quality and consistency of instruction in the classroom to reflect the instructional principles derived from scientifically-based research in reading; (b) *Reading First* schools must improve the use of information obtained from early reading assessments so that struggling readers can be identified and provided additional instruction in a timely manner; and (c) *Reading First* schools must establish procedures to provide those struggling readers with intensive interventions to supplement the instruction received in the regular classroom (Florida Department of Education, 2002).

As defined by the FDOE (2002), an effective reading program provides systematic, high quality instruction that focuses on the five (5) components of reading instruction: phonemic awareness, phonics, fluency, comprehension and vocabulary. It should also incorporate the appropriate use of three (3) important types of assessment to guide reading instruction (screening, diagnostic and progress monitoring). Initial Instruction must be grounded in scientifically-based reading research and aligned with the Florida Sunshine State Standards for reading. In addition, Immediate Intensive Intervention must be used to help struggling readers make adequate progress in learning to read.

Oral Reading Fluency

This measure has consistently been shown to be an effective predictor of reading comprehension, the overall goal of reading instruction (Fuchs, Fuchs, Hosp & Jenkins, 2001; Good, Simmons & Kame'enui, 2001). Oral Reading Fluency is a measure that assesses fluency with text, the ability to translate letters-to-sounds-to-words fluently and effortlessly (LeBerge & Samuels, 1974). The fluent reader is one whose decoding processes are automatic, requiring no conscious attention (Juel, 1991). Such capacity then

enables readers to allocate their attention to the comprehension and meaning of the text.

Limitations to the Study

Principals selected for this study provided self-report responses to the items contained in the PIQ. As such, responses must be accepted as valid without confirmation of accuracy.

Delimitations to the Study

1. Oral reading fluency (ORF) assessment scores are the only outcome measures selected for this study. Although there is a base of research to support this selection, there exists the potential that relevant relationships between the program implementation behaviors of principals and student achievement in other aspects of reading remain undetected.

2. Only the ORF assessments of first grade students in Florida *Reading First* schools were analyzed to determine relationships between the program implementation behaviors of principals and student achievement in reading. This design limits the generalizability of the results across elementary grade levels.

3. It is important to note that the validation of the Principal Implementation Questionnaire (PIQ) was an inductive process, meaning that the revision and development activities were conducted on survey constructs, associated domains and many individual items that were previously constructed. As such, the final version of the survey instrument used in this study is somewhat limited by the parameters set by the previous version of the survey.

4. Principal behaviors associated with the implementation of effective reading programs are, inherently, a level removed from actual student instruction. It is generally accepted by researchers that a vast majority of the school-level student achievement variance is accounted for at the classroom level and is dependent on teacher implementation of effective reading programs. That being said, the central purpose of this study was to determine the principal-level effects on reading achievement, without regard for classroom-level variables.

CHAPTER 2

REVIEW OF THE LITERATURE

During the past 30 years there has been an increasing focus on students in public schools who are not achieving academically. This focus has led many to search for reasons for the lack of achievement and, from the theoretical basis of these reasons, design educational programs to better assist these students and their teachers. Effective educational leadership makes a difference in improving learning; there is nothing new or especially controversial about this idea. What's far less clear, even after several decades of school improvement efforts, is just how leadership matters and how important those effects are in promoting the learning of all children. For example, the available evidence about the size and nature of the effects of successful leadership on student learning demonstrates that leadership is second only to classroom instruction among all school-related factors that contribute to what students learn at school (Leithwood, Louis, Anderson & Wahlstrom, 2004). Much of the research on the effects of leadership on student learning is unclear, at best. This is due, in part, to the fact that the methodologies employed by many of these studies significantly underestimate actual effects. Some researchers have shown that leadership effects (both direct and indirect) account for up to one-quarter of total school-level effects (Hallinger & Heck, 1996, 1999; Leithwood & Jantzi, 2000).

The following review of the literature will begin with a focus on out-of-school variables that effect student achievement. Then, research related to in-school factors that effect student achievement will be reviewed, specifically including the concept of the principal as the instructional leader. This review will conclude with (a) a discussion of emerging issues in investigations of the relationship between the instructional leaders (i.e., principals) and student achievement and (b) conclusions and implications for future research.

Exogenous Variables of Student Achievement

Throughout the body of literature on the subject, there have been many variables cited as contributing to student achievement. The body of empirical research predating

the 1990s was generally focused on the impact of exogenous (out-of-school) variables on student achievement. Early research in this area focused primarily on the background of the student and its potential effect on student performance. Some early studies hypothesized that students from more affluent backgrounds would perform better in school than those of less affluent backgrounds, citing the socioeconomic status (SES) of individual students before they entered school as the primary indicator of academic success.

Student-level characteristics including family demographics, the innate intellectual ability of the student and the student's motivation for learning, are consistently shown to have the strongest impact on student academic achievement. Thus, school-level variables including school leadership, teacher qualifications, instructional practices and related services explain a much smaller portion of the total student-level variance in academic performance (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld & York, 1966; Jencks et al., 1972; van de Grift & Houtveen, 1999). This view guides findings that show endogenous (in-school) variables as weak predictors of student achievement. From this, many researchers support the idea that the provision of additional school-level resources will result in insignificant changes in individual student achievement (Coleman et al., 1966; Hanushek, 1986, 1989, 1996, 1997, 1999; Jencks, 1972; van de Grift & Houtveen, 1999).

Establishing a clear relationship between school-level resources and student achievement is a challenging goal. One of the first attempts to do this was *Project Talent* (Flanagan, Tiedeman, Clemans & Wise, 1964), a research project commissioned by the U.S. Department of Education. This study was conducted using high school students and investigated the relationship between the interests, talents and aptitudes of students and student achievement. The study was designed to (a) determine the effects of motivation on student achievement, (b) identify variables that affected the vocational choices of students, (c) identify possible predictors of creativity and productivity among students, (d) determine the effectiveness of different educational experiences on student achievement and (e) investigate the methods for identifying academic potential in students. While it did not seek to directly tie resources to student achievement outcomes, the Project Talent final report provided extensive information on the interests, aptitudes

and innate talents of students, which guided much of the subsequent literature on school-level factors effecting student achievement.

One of the most widely cited and influential studies relating resources to achievement was *Equality of Educational Opportunity*, or the “Coleman Report,” (Coleman et al., 1966). Coleman studied approximately 600,000 children at 4,000 schools throughout the United States. This was the first research design to focus on educational outputs in relation to educational inputs, such as per-pupil expenditures and teacher characteristics. Multiple regression analysis was used to estimate and predict the relationship between certain school resources and student achievement, while controlling for demographic and other background characteristics.

In this study, the school outcome of interest was student learning as measured by student achievement tests. Several characteristics of schools, including teacher-student ratios, were shown not to be key predictors of student success when SES was a considered variable (Coleman et al., 1966). This study also concluded that factors external to the school were more influential than in-school variables in predicting achievement.

Researchers have continued to find that schools are similar in the way they relate to the achievement of students when SES is taken into account (Hallinger & Heck, 1999; Taylor & Tashakkori, 1995). Student characteristics have persistently been shown to have the strongest effect on student achievement. In fact, when student-level variables are controlled for, school-level variables are consistently shown to have a small impact on student learning (Coleman et al., 1966; Hanushek, 1986, 1989, 1996, 1997, 1999; Jencks, 1972, 1977). Furthermore, Coleman (1966) suggested that the differences in schools attended by white students and black students were (a) overstated, (b) made little difference in actual student achievement even if they did exist and (c) that most children attended schools where they were the majority race. In-school variables such as teacher training, teacher salaries and type of curriculum used were relatively equal (Coleman et al., 1966).

A student’s peer group, the values that they are exposed to and a student’s satisfaction or dissatisfaction with the school, are related to the overall population of students within their school. Jencks (1972) reanalyzed Coleman's data and also concluded

that the achievement of lower class students, regardless of race, was highly related to the socioeconomic level of their classmates. This finding is in agreement with Coleman's argument (1966) that school social class has an impact on individual student academic achievement independent of the factors related to a student's family.

Jencks (1972) concluded that family SES correlated at 0.35 with student academic performance. That is, the academic performance of students in the top 20% in terms of SES should score approximately 13 to 15 points higher on tests of academic performance than students in the bottom 20%. Jencks also found that SES disparities are greatest for general information and verbal ability scores.

Endogenous Variables of Student Achievement

Much of the research relating school resources (inputs) to student achievement (outputs) has been rooted in economics. Schultz (1961) studied the advantages of investing in education to improve agricultural production. He suggested that investing in the educational attainment of workers (human capital) would produce a more robust U.S. economy than investing in physical capital, such as factories and machinery.

Becker (1964) suggested that human capital could be developed and accumulated through investments in education and training. Further, the accumulation of human capital would build capacity for increased production in most areas of the economy. As educational institutions are the main resource for providing these types of opportunities, researchers began to look for school-level factors that could be improved to maximize such resources. The research of Shultz (1961) and Becker (1964) initiated a body of education production function research which focused primarily on school spending and the resources that educational funds can purchase. To the present, education production function research remains one of the primary ways to view the relationship between educational inputs and outputs.

Education Production Function Research

The establishment of a link between human capital and education created the need to explain the relationship between school-level resources and educational outputs. Previous research had shown that variables based on individual student characteristics (exogenous variables) accounted for most of the variability in student achievement. This

knowledge led to the creation of an education production function equation, which controlled for student demographic variables (Belfield, 2000). This equation controlled for variables not accounted for at the school level and allowed researchers to isolate the effects of in-school resources on educational output (student achievement). Researchers were then able to measure the actual impact of school-level variables on student achievement.

Seeking to determine the importance of various resources in education, much of the research in this area has focused on variables that can be purchased, thereby controlling the amounts available to schools. This methodology follows basic economic principles that suggest that an increase in relevant resources should result in an increase in educational production. In this light, education production function methodology depicts schools as factories “producing” student achievement (Greenwald, Hedges & Laine, 1996).

Much of the interest in this area of research was because per-pupil expenditures nationwide doubled from the late 1960s to the early 1990s (adjusting for inflation) with little, if any, increase in student achievement (Darling-Hammond, 1999; Greenwald, Hedges & Laine, 1996). In fact, one of the most visible and cited outcome measures—SAT scores—declined during this time period and scores on the National Assessment of Education Progress (NAEP) showed no significant increases. This and other information on the topic led researchers such as Hanushek (1986, 1989, 1996, 1997, 1999) to conclude that additional financial resources have no appreciable effect on achievement.

Diverging from SES-based approaches, Hanushek (1986, 1989, 1996, 1997, 1999) conducted a meta-analytic synthesis of a selected group of education production studies. This line of research concluded that the data reported in the body of education production literature failed to establish a direct relationship between school resources and student achievement. Hanushek’s conclusions, while controversial, garnered substantial support and was accepted by many in the areas of education, research and policy (Greenwald, Hedges & Laine, 1996).

Hanushek’s work in this area is a much debated topic and continues to influence research on resources and student achievement and related policy decisions. Hanushek (1986) found that unique individual traits of classroom teachers had more impact on

student achievement than the typically cited variables of school quality, such as class size and per-pupil expenditure. In subsequent meta-analyses of educational spending research, Hanushek (1989, 1991, 1997) analyzed studies that correlated school-level spending with student achievement. This methodology examined 38 studies conducted from the 1960s through the 1980s to investigate the possible relationships between educational inputs and student outcomes. A total of 187 production function analyses were identified measuring seven variables, including (a) per-pupil expenditures, (b) teacher experience, (c) the educational attainment of teachers, (d) teacher salaries, (e) student/teacher ratios, (f) school administration (i.e., principal experience) and (g) factors associated with school facilities. As a common theme running through this line of inquiry, Hanushek (1989) found that there was a very small percentage of positive and statistically significant correlations between economic inputs and student achievement outcomes for each of the seven variables. This finding prompted Hanushek to state that there was “no strong or systematic relationship between school expenditures and student performance” (p. 47). School resources, therefore, had an insignificant impact on student achievement.

Hedges, Laine and Greenwald (1994) replicated Hanushek’s study, reaching different conclusions. The authors noted that although a majority of the relationships were statistically insignificant, when examined in isolation, over 70% of the relationships were, in fact, positive and of sufficient magnitude. Their analysis concluded that school resources were directly related to student achievement in ways that are vital. In response to Hanushek, they stated that “relying on the data most often used to deny that resources are related to achievement, we find that money does matter after all” (p.14).

Later literature reviews also challenged Hanushek’s interpretation of the student achievement research to date. Much of the criticism focused on Hanushek’s methodology for selecting and weighting studies used in his analyses. For example, both Hedges et al. (1994) and Krueger (2000) used more refined statistical methodologies to compare the studies that Hanushek used in his previous work. In each case, the researchers determined that educational resources were related to student achievement in positive and significant ways.

Hedges et al. (1994) used a meta-analytic technique that weighted the individual studies on a proportional basis, not equally as Hanushek had done. This study used

statistical formulas for student achievement variables to reduce the overall measurement bias within the population of studies. As a result, this analysis showed that most of the student achievement variables that showed statistically insignificant results in Hanushek's analyses actually displayed positive relationships to student achievement. Pupil-teacher ratios, the amount of education received by teachers and the amount of teacher salary each produced weak positive relationships, while the amount of teaching experience held by classroom teachers as well as per-pupil expenditures showed stronger positive relationships (Grissmer, Flanagan, Kawata & Williamson, 2000).

Krueger (2000) also analyzed the sample of existing student achievement literature used by Hanushek (1997) and refuted his conclusions. This study challenged the selection and weighting methodology of Hanushek, as well. Krueger (2000) argued that the statistical power of many of the studies used by Hanushek were diminished by the inclusion of multiple subgroups, resulting in insignificant or negative relationships (Grissmer et al., 2000). Krueger (2000) also criticized Hanushek's method that treated each of the measures within a study as an individual estimate of the relationship between school resources and student achievement.

While the opposing views of these studies demonstrated the lack of comprehensive understanding and agreement in this area, there have been additional studies conducted over the past decade that establish a direct link between resources and student achievement (Biddle, 1997; Ellinger, Wright & Hirlinger, 1995; Wenglinsky, 1997). For example, Ellinger, Wright and Hirlinger (1995) investigated the relationship between fiscal inputs and academic achievement among 11th grade students in Oklahoma school districts. The study concluded that per-pupil revenues within individual schools were positively related to increased achievement. Similarly, Fortune and O'Neill (1994) found positive relationships between instructional expenditures and student achievement within school districts in Ohio and Missouri.

Hartman (1994) also examined whether higher levels of spending resulted in higher student achievement. In this study, Hartman assessed whether there were differences in district spending among school districts and whether or not these differences were in areas known to be related to student achievement. The study concluded that more affluent school districts used resources toward (a) reducing

student/teacher ratios, (b) employing more teachers, (c) employing more qualified teachers, (d) increasing teachers' salaries and (e) employing more administrative and support staff. In addition, using data from the National Assessment of Educational Progress (NAEP), Wenglinsky (1997) found that higher per-pupil expenditures, resulting in smaller class sizes, had a positive impact on national 8th grade mathematics achievement.

Overall, this line of economics-based research sheds light on the relationship between school-level funding and the in-school resources that those dollars provide and student performance. While there is a considerable amount of residual controversy on this topic, there exists much evidence that the effective use of school-level resources can produce significant gains in student performance.

Effective Schools Research

The *Coleman Report* (Coleman et al., 1966) suggested that individual schools were not important in predicting student achievement. This produced a backlash of research which would become the base for the effective schools movement. Edmonds (1979) authored what is widely regarded as the seminal work of the effective school movement. Later, he published a paper entitled *Programs of School Improvement: An Overview*, which identified the correlates of effective schools (Edmonds, 1982). In this paper, Edmonds stated that all effective schools share several things in common, including (a) principal leadership that supports and facilitates quality instruction, (b) an instructional focus that is clear and widely communicated, (c) a safe instructional environment conducive to teaching and learning, (d) high learning expectations for all students and (e) the use of student achievement goals as the basis of evaluation.

The results of the original effective schools research and those of later studies on the topic have identified elements that comprise a successful school (Edmonds, 1979; Brookover & Lezotte, 1982; Lezotte, 2001). Effective schools research gained prominence in the 1980s with the release of the correlates of effective schools. The correlates include (a) instructional leadership, (b) clear and focused mission, (c) safe and orderly environment, (d) climate of high expectations, (e) frequent monitoring of student progress, (f) positive home-school relations, and (g) opportunity to learn and student time

on task (Lezzote, 2001). These elements are now widely recognized as necessary for school success.

As the correlates have impacted educational leadership research profoundly in recent decades, a more comprehensive discussion of their content is warranted. The following is essential detail meant to clarify the components of the correlates of effective schools (Levine & Lezotte, 1990; Lezzote, 2001).

Instructional Leadership. An effective school has a principal that serves as an instructional leader and effectively communicates the mission of the school to all stakeholders. The principal also uses effective instructional practices to guide and support an instructional program that is focused on student achievement.

Clear and Focused Mission. An effective school has a clearly articulated mission, or statement of purpose, that communicates and instills a sense of shared responsibility among stakeholders to pursue the goals and priorities of the school.

Safe and Orderly Environment. An effective school displays an environment where all stakeholders feel safe from harm and are free to pursue appropriate academic and social goals. The school climate promotes collaboration and is conducive to effective teaching and learning. Also, desirable student behaviors are consistently communicated and expectations are clear.

Climate of High Expectations. An effective school promotes a climate of high student expectations where all students are encouraged to meet standards set forth by the curriculum. Staff in an effective school feel that they can teach all students to meet at least minimum performance requirements. In order to meet these high expectations, the school is focused on student learning not instructional practice.

Frequent Monitoring of Student Progress. An effective school has processes in place to frequently monitor pupil progress in essential areas, using a variety of assessment procedures that emphasize authentic assessment in curriculum mastery. In addition, staff in an effective school effectively use data gathered through these monitoring processes to inform and guide instructional practice to the benefit of the students and the school.

Positive Home-School Relations. An effective school communicates the mission of the school to parents and guardians and allows these stakeholders to contribute to the

school community in important ways. Forming partnerships with the parents and the community enables all stakeholders to have the same goals and expectations.

Opportunity to Learn and Student Time on Task. An effective school promotes the use of a majority of the school day for teacher-driven classroom instruction in essential curricular areas, abandoning less important content. Effective instruction is guided by an integrated, interdisciplinary curriculum.

In the body of effective schools research, Levine and Lezotte (1990) authored one of the more comprehensive syntheses, with a focus on research studies conducted between 1985 and 1990. This research focused on (a) the value of the effective school model in school improvement efforts, (b) the relationship between research and basic findings and (c) the alignment of school practices with the related research on effective school practices. The findings of this study suggest that the correlates should be considered as prerequisites for school success rather than guarantees of success.

Additional Endogenous Factors of Student Achievement

Research has demonstrated that potential in-school factors related to student achievement are numerous. The list of school-level correlates of student achievement includes factors such as school and class size, instructional quality, teacher qualifications, school climate and instructional leadership. However, unlike the social background factors, these features are endogenous and to varying extents under control of the principals and teachers in the school.

School/Class Size. Research evidence is mounting to support the argument that student achievement is lower in bigger schools. Larger school size has been shown to be associated with reduced academic engagement among eighth-grade students (Lee & Smith, 1993). Lee & Smith (1995) then concluded that larger schools display reduced student engagement among high school students. Lindsay (1982) had similarly identified students from larger schools as displaying lower levels of satisfaction with school experiences. Also, Pittman & Haughwout's (1987) research showed that students in larger schools displayed lower school attendance and higher dropout rates than their counterparts in smaller schools. It has also been argued that students in smaller schools have a greater sense of belonging. Further, school personnel in these schools are more likely to know students by name and to identify at risk students (Cotton, 1996). Another

arguable advantage to smaller schools is the finding that teachers and administrators may be more positive about their roles because smaller schools provide additional flexibility in practice and may afford these positions more opportunities to make a difference for individual students (Cotton, 1996). Such findings have resulted in calls for larger schools to be sub-divided into smaller, more intimate learning communities (Boyer, 1995; Oxley, 1994).

It is now commonly accepted in practice that smaller classes are better environments for learning, allowing for more individual instructional time and attention (Finn & Achilles, 1990; Krueger, 2000; Krueger & Whitmore, 2001; Mosteller, 1995). In an investigation of the impact of smaller classes on instructional environment, Blatchford & Mortimore (1998) found that in relation to smaller classes, (a) teachers appeared more able to maintain control and a level of noise that aided students in staying on task and completing assignments, (b) teachers were more able to offer timely feedback that was tailored to meet individual student needs and (c) students did not have to wait for teacher feedback and attention, thereby becoming discouraged or distracted. There is a mounting body of research that supports this position (Blatchford & Mortimore, 1994, 1999; Finn & Achilles, 1990).

Instructional Quality. When considering school-level effects on student achievement, enhanced quality of instruction has consistently been shown to relate positively to academic performance (Scheerens & Bosker, 1997; Stallings & Mohlman, 1981; Wang, Haertel & Walberg, 1993). For example, research evidence suggests that students tend to learn more when their teachers use quality instructional techniques and appropriate curricula. Further, teachers that use active teaching strategies are more effective (Brophy & Good, 1970, 1986; Rowan, Chiang & Miller, 1997), as are teachers that have access to more opportunities to learn effective pedagogical skills and strategies (Cohen & Hill, 1997).

The importance of effective instructional practices in regard to student achievement has been documented in numerous studies. For example, Sanders and Horn (1998) looked at the levels of student achievement among a sample of students of teachers categorized from “most effective” to “least effective” in their teaching practices. In this study, the researchers found large differences in student achievement between the

students taught by “most effective” teachers and those instructed in the classrooms of “least effective” teachers. The researchers found “that race, socioeconomic level, class size and classroom heterogeneity are poor predictors of student academic growth. Rather, the effectiveness of the teacher is the major determinant of student academic progress” (Sanders & Horn, 1998, p. 247).

In an attempt to clarify the components of effective instruction, Waters, Marzano and McNulty (2003) conducted a meta-analysis review of effective instruction and identified nine instructional strategies that are associated with increased student achievement. These included (1) identifying similarities and differences, (2) summarizing and note taking, (3) reinforcing effort and providing recognition, (4) homework and practice, (5) nonlinguistic representations, (6) cooperative learning, (7) setting goals and providing feedback, (8) generating and testing hypotheses and (9) activating prior knowledge. They contend that by integrating these research-based instructional strategies into their current classroom practice, teachers can improve student understanding of instructional content and performance on outcome assessments.

Teacher Qualifications. Another endogenous factor that has been shown to be related to student achievement is the formal qualifications of classroom teachers. This variable takes into account the preparation and experience of the teacher. As such, more qualified teachers have more training and classroom experience, making them generally better equipped to manage the classroom and providing a greater knowledge of instructional content and pedagogical techniques (Darling-Hammond, 1999, 2000). Further, student achievement improves with higher-levels of teacher qualifications (Darling-Hammond & Youngs, 2002; Monk, 1994; Rowan et al., 1997). Therefore, it is easier for experienced, qualified teachers to create a more effective learning experience for students, especially low-performing students.

In a nationwide study of teacher effectiveness, Darling-Hammond (1999) examined how teacher qualifications and other school inputs are associated with student achievement across the 50 states. The findings of both the qualitative and quantitative analyses suggest that the quality of teachers may be related to improvements in student performance. Other large-scale analyses (i.e., Rowan et al., 1997) also suggested that a

significant amount of the variance in student achievement among elementary school classrooms can be explained by factors associated with teacher qualifications.

School Climate. School climate refers to the atmosphere created by staff and student behavior in the school. School climate is a multi-dimensional concept that encompasses aspects of school organization, staffing, the student body and the community. It has been determined that a positive school climate can yield positive educational outcomes for students and school staff. Conversely, a negative school climate can be a deterrent to effective instruction and student learning (Freiberg, 1998; Johnson & Johnson, 1993, 1997; Kuperminc, Leadbeater & Blatt, 2001; Kuperminc, Leadbeater & Blatt, 2001; Manning & Saddlemire, 1996). Peterson and Skiba (2001) also contend that principals are the key to shaping a positive school culture. In this light, the role of school principal is demonstrated to have an impact on individual student learning.

Other school climate research suggests that a positive school climate can increase student achievement levels and reduce disruptive behavior (McEvoy & Welker, 2000; Samdal, Nutbeam, Wold & Kannas, 1998). Research has also shown that students that have a positive view of their school's climate are more motivated and display higher achievement than those that are not as satisfied with the climate of their school (Hoy, Smith & Sweetland, 2002; Rutter, Maughan, Mortimore, Ousten & Smith, 1979; Mortimore, Sammons, Stoll, Lewis & Ecob, 1988; Voelkl, 1995).

The Principal as Instructional Leader

As the role of the principal has evolved, the concept of instructional leadership emerged as a way to categorize the activities and responsibilities of principals in relation to classroom instruction (Deal & Peterson, 1990). Practitioners and researchers have developed many differing definitions of instructional leadership since that time. The National Association of Elementary School Principals (2001) defined instructional leadership as "leading learning communities". This definition views principals as facilitators, guiding and encouraging an educational environment in which administrators and teachers work collaboratively to diagnose and solve the problems facing their schools.

Blase and Blase (2000) defined instructional leadership in a series of seven principal behaviors: (a) making suggestions, (b) giving feedback, (c) modeling effective instruction, (d) soliciting opinions, (e) supporting collaboration, (f) providing professional development opportunities and (g) giving praise for effective teaching. Effective principals have also been said to display eight common effective traits: (a) recognizing teaching and learning as the main business of the school, (b) communicating the school's mission clearly and consistently to all stakeholders, (c) fostering standards for teaching and learning that are high and attainable, (d) providing clear goals and monitoring the progress of students toward meeting them, (e) spending time in classrooms and listening to teachers, (f) promoting an atmosphere of trust and sharing, (g) building an effective staff and making professional development a top priority and (h) not tolerating ineffective teachers (Education Week, 1998). Moreover, as previously detailed, studies of effective schools have identified five instructional leadership priorities of effective principals (a) defining and communicating the school's educational mission, (b) managing curriculum and instruction, (c) supporting and supervising teaching, (d) monitoring student progress and (e) promoting a learning climate (Bateman & Bateman, 2001; Blase & Kirby, 1992).

In recent decades, the importance of effective instructional leadership on school performance has been well documented in the literature (Gates, Ross & Brewer, 2001; Leithwood, 1988; Purkey & Smith, 1983; Senge, 1990). While there are numerous constructs by which the components of effective leadership are defined, there is also a great deal of similarity among them. A consensus on the definition of effective school leadership is far from being reached; however, there are several identifiers that are commonly held as being critical factors of effective leadership.

Safe and Orderly Environment

One of the most fundamental responsibilities of a school principal is to provide a safe and orderly educational environment that allows for effective teaching and learning. Researchers have identified several factors of a safe and orderly environment that can be affected by principal behavior including (a) the setting and communication of behavioral standards, (b) implementing effective processes to ensure that behavioral policies are applied consistently for all students, (c) assuring that discipline is used consistently and

fairly and (d) dispersing the responsibility for discipline throughout the school, among others (Cotton, 2003; Marcoulides & Heck, 1993; Leitner, 1994; Sammons, Hillman & Mortimore, 1995; Scheurich, 1998).

Mission and Vision

The importance of a clear mission and vision to successful schools has been frequently supported in the literature (Cotton, 2003; Johnson & Asera, 1999; Louis & Miles, 1992; National Commission on Education, 1995; Scheurich, 1998). Researchers have identified transformational leadership, which is focused on improving the overall culture and organization of a school, as being closely related to clear understandings of vision and mission (Hallinger, 2000; Hallinger, 2003; Hallinger & Heck, 2002; Murphy & Louis, 1994; Leithwood & Jantzi, 2000; Fullan, 2003). Similarly, purpose-driven leadership is often related to schools that are considered effective (Rutter et al., 1979; Mortimore et al., 1988; Teddlie & Stringfield, 1993; Sammons, Thomas & Mortimore, 1997).

Stakeholder Involvement

Another component of effective leadership is the principal's ability to garner outside resources toward the improvement of the school. Researchers consistently cite community/stakeholder involvement as related to high-achieving schools (Marcoulides & Heck, 1993; Johnson & Asera, 1999; Sammons, Hillman & Mortimore, 1995; Scheurich, 1998; Yap & Enoki, 1995). To this end, effective principals have been shown to (a) build the leadership capacity of teachers and staff, (b) encourage team learning focused on schoolwide goals, (c) use organizational flexibility to enhance effectiveness and (d) distribute leadership responsibilities throughout the school (Rea, McLaughlin & Walther-Thomas, 2002; Yap & Enoki, 1995).

Monitoring School Progress

Personal monitoring of school progress by the principal has been shown as a predictor of school effectiveness in most studies where it has been included as a variable (Levine & Lezotte, 1990). It is generally held that effective principals routinely visit classrooms, participate in team-level meetings and pay close attention to student performance within their school (Elmore, 2000; Fink & Resnick, 2001; Gullat & Lofton, 1996; McCallum, 1999; Sammons et al., 1995). Further, it has been argued that personal

interactions are the best way for a principal to affect positive change within a school (Deal & Paterson, 1990).

Murphy (1990) found that effective principals utilized several monitoring strategies, including (a) using assessment to inform instruction; (b) communicating information on student data to all stakeholders; and (c) constantly evaluating the instructional quality and academic progress of the school. Effective principals have also been shown to routinely use school- and student-level data to guide programmatic and instructional decisions (Leithwood & Jantzi, 2000; Spillane, Halverson & Diamond, 2001; Stringfield, 1995; Reynolds & Stringfield, 1996).

Instructional Focus

One of the key responsibilities of an instructional leader is to maintain a school-wide focus on critical instructional areas. Principals of effective schools have been shown to take personal interest and responsibility for instructional matters (Klingner, Arguelles, Hughes & Vaughn, 2001; Leithwood & Jantzi, 2000; Waters et al., 2003). Johnson and Asera (1999) found that high-performing principals created opportunities for teachers to plan and work together regarding instructional issues. In fact, the school schedule was planned around the key instructional needs of the school in an effort to protect critical instructional time (Cotton, 2000; Evans & Teddlie, 1995; Johnson & Asera, 1999). Heck (1993) identified the principal's willingness and ability to promote meaningful discussion around instructional issues as central to their effectiveness as a school leader.

Murphy (1990) stated that there are three areas where leadership in instructional focus is most important: (a) creating focused school goals and communicating them to stakeholders; (b) managing the instructional environment by frequent monitoring of instructional processes; and (c) promoting an academic learning climate by maintaining high expectations, providing sufficient instructional resources and ensuring adequate professional development opportunities for teachers.

High Expectations for Student Performance

Research has also pointed to the performance expectations held by the principal as an important aspect of effective schools. Consistently communicating expectations for high performance has been linked by researchers to positive results in school and student achievement (Cheng, 1994; Davis, 1998; Gullat & Lofton, 1996; Leithwood & Jantzi,

2000; Waters et al., 2003). Teddlie and Reynolds (2000) found that high-performing principals monitor classroom-level expectations to ensure alignment with the high expectations of the school. They further suggest that effective principals: (a) expect staff to work at understanding school conditions and issues before they start work, (b) expect high levels of participation in professional development activities, (c) expect high-quality instructional practice, (d) expect staff to prioritize student achievement as the primary goal, and (e) expect staff focus time management towards instructional priorities (Teddlie & Reynolds, 2000).

Professional Development

Research provides extensive support for the idea that much of a principal's success comes through the professional development opportunities that they provide for their staff, especially those in instructional positions (Deal & Peterson, 1990; DuFour & Berkey, 1995; Leitner, 1994; Levine & Lezotte, 1990; Peterson, Gok & Warren, 1995; Wagstaff, Melton, Lawless & Combs, 1998). Effective principals have also been shown to participate in professional development activities in order to gain understanding of classroom practices (Wagstaff et al., 1998). An additional practice associated with professional development that has been cited as being displayed by effective principals is the ability to acquire professional development resources for their school. This includes time for training, funding to pay for training as well as professional development materials. (Bamburg & Andrews, 1991).

Principal Behaviors and Student Achievement

There is little controversy over whether educators and educational researchers think that school principals make a positive impact on school performance. Indeed, there is ample evidence in the body of research and in educational practice to confirm that the school principal is regarded as critical to school success and student achievement. In particular, several decades of research on the topic has resulted in a body of knowledge that details the positive relationships between the practice of school principals and student academic achievement (Cotton, 2003). The following review of the related

literature seeks to synthesize the major studies on the topic over the past 30 years, with a special focus on research covering the last 15 years.

Despite the large body of literature on educational leadership, the causal relationship between principal behaviors and student achievement remains unclear, at best (Hallinger, Bickman & Davis, 1996; Witziers, Bosker & Kruger, 2003). Some studies in this area investigate the relationship between school-level variables and student achievement, yet fail to bring specific principal behaviors into the model. Examples of this type of study includes those focused on school mission (Bossert, 1988), school culture (Deal & Peterson, 1999), organizational considerations such as school size (Lee & Loeb, 2000; Lee & Smith, 1995), or the placement of highly-qualified teachers in school classrooms (Ingersoll, 1996). A second type of study investigates the principal's role in shaping the educational environment, but does not use student achievement as a dependent variable, (e.g., Sanders' and Harvey's 2002 study on the impact of community collaboration in schools). The lack of consensus regarding the components of leadership has led to the proliferation of these types of "incomplete" methodologies.

In addition, much of the research conducted on the potential relationship between principal behaviors and student achievement has consisted of basic statistical methodologies relying on simple, bivariate statistical models (Hallinger & Leithwood, 1994; Hallinger et al., 1996; Witziers et al., 2003). This type of research generally seeks to establish a direct link between the actions of principals and the achievement of students. Contemporary scholars have, for the most part, discarded the view that principals have a direct effect on student learning and have, instead, adopted the notion that principals have an impact on student achievement through their interactions with teacher and their efforts in shaping and managing the educational environment (Hallinger & Leithwood, 1994; Heck, 1993; Hallinger et al., 1996; Witziers et al., 2003).

Hallinger et al. (1996) also suggested that because previous research in this area has been weakened by the use of oversimplified methodological approaches, forthcoming research designs will be strengthened if they include (a) sufficient sample sizes, (b) theoretically defensible models, (c) reliable data collection instruments and (d) sophisticated data analysis tools. The use of these elements should enable researchers to better capture the indirect effects of principal behaviors on student achievement

(Hallinger et al., 1996; Witziers et al., 2003).

There are, however, some effective models of research focusing on leadership that include the behaviors of principals and their effects on individual student achievement in the research design. Hallinger and Heck (1996) identified approximately forty such studies conducted between 1980 and 1995. They included studies that investigated principal effects on student achievement measures, whether those effects were direct or indirect. The study concluded that research incorporating sophisticated modeling methods showed that the effects of school-level leadership on individual student achievement was generally small. However, these results appeared to be educationally significant in relation to the small proportion of student-level variance that can be explained outside of exogenous (student-level) variables. Leadership effects were shown to explain only up to 5% of the total variance; however, this translated into approximately 25% of the total student-level variance explained by school-level variables (Hallinger and Heck, 1996).

Quantitative studies that included leadership and student achievement in their models were also reviewed by Scheerens and Bosker (1997) with similar conclusions. Using a modeling methodology to investigate the possible relationships between principal behaviors and student achievement, the researchers included school leadership as one possible indicator of school effectiveness. They found that variables more closely associated with the student (classroom-level variables) produced more robust relationships than did variables at the school-level (i.e., leadership). In a meta-analysis of additional research on the subject, Scheerens and Bosker (1997) found small, but statistically significant principal effects on student achievement. Further, the study showed that controlling for various contextual variables led to stronger principal effect sizes.

In a meta-analysis of 70 contemporary classroom and leadership studies, Waters et al. (2003) investigated whether the quality of leadership has a significant relationship to student achievement and what specific leadership responsibilities and practices have the greatest impact. They concluded that school leadership is an important variable, as it correlates positively with student achievement. The researchers identified 21 key areas of leadership that correlate positively with student achievement. These key areas are: (1) culture; (2) order; (3) discipline; (4) resources; (5) knowledge of curriculum, instruction

and assessment; (6) focus; (7) visibility; (9) contingent rewards; (10) communication; (11) outreach; (12) input; (13) affirmation; (14) relationship; (15) change agent role; (16) optimizer role; (17) ideals and beliefs; (18) monitoring and evaluation; (19) flexibility; (20) situational awareness; and (21) intellectual stimulation.

Waters et al. (2003) also stated that effective principals have a comprehensive knowledge of leadership strategies and have developed an awareness of when to use them. Further, they understand how to balance school culture, the student population and the community in order to promote increased student achievement. Two variables were identified as determining factors in whether principal effects on student achievement are positive or negative: (a) correctly identifying the focus for improvement and (b) understanding how closely the proposed change matches existing values norms and values.

Conceptual and methodological challenges notwithstanding, previous research has identified a measurable impact of effective principal leadership on individual student achievement. Principal effects have been shown to be primarily indirect, as they are typically mediated by other variables more proximal to the student level. Moreover, student achievement effect sizes in relation to principal leadership have proven to be small. These proportions of student-level variance are, however, practically (and statistically) significant. The importance of these findings is amplified when considered in light of the relatively small proportion of individual student achievement variance that can be attributed to endogenous variables.

CHAPTER 3

METHODS

Purpose and Research Questions

The purpose of this study was to determine the relationship between the principal's role in the implementation of reading programs and student achievement in reading. Using archival data on all Florida schools receiving *Reading First* grants in the 2004-2005 academic year, hierarchical growth curve modeling was used to determine if there is a significant relationship between principal practices in the implementation of effective reading programs and growth in student reading achievement.

The following research questions and relative hypotheses were addressed in this study:

1. Is there a significant relationship between defined elements of principals' implementation of effective reading programs and growth in student reading achievement in the total population of first grade students identified during the 2004-2005 academic year?

Null Hypothesis (H_{01}): There is no statistically significant difference in the growth trajectories of student reading achievement of students in schools where the principal displays high levels of implementation of effective reading programs versus those displaying lower levels.

Research Hypothesis (H_{a1}): There is a statistically significant difference in the growth trajectories of student reading achievement of students in schools where the principal displays high levels of implementation of effective reading programs versus those displaying lower levels.

Since the Principal Implementation Questionnaire (PIQ) was designed to measure five distinct aspects of reading program implementation, five sub-hypotheses were considered—in relation to each research question—to identify specific effects of each of the PIQ domains on the oral reading fluency of the total population of students.

1.1 There is no statistically significant effect of principal behaviors in the implementation of effective reading programs on the reading achievement of students in

relation to the PIQ Core Reading Program domain.

1.2 There is no statistically significant effect of principal behaviors in the implementation of effective reading programs on the reading achievement of students in relation to the PIQ Leadership domain.

1.3 There is no statistically significant effect of principal behaviors in the implementation of effective reading programs on the reading achievement of students in relation to the PIQ Professional Development domain.

1.4 There is no statistically significant effect of principal behaviors in the implementation of effective reading programs on the reading achievement of students in relation to the PIQ Assessment domain.

1.5 There is no statistically significant effect of principal behaviors in the implementation of effective reading programs on the reading achievement of students in relation to the PIQ Intervention domain.

2. Is there a significant difference between the reading achievement of first grade students of low socio-economic status (SES) and students of higher socio-economic status in relation to principal effects?

Null Hypothesis (H_0): There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to principal effects.

Research Hypothesis (H_a): There is a statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to principal effects.

2.1 There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to the PIQ Core Reading Program domain.

2.2 There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to the PIQ Leadership domain.

2.3 There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to the PIQ Professional Development domain.

2.4 There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to the PIQ Assessment domain.

2.5 There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to the PIQ Intervention domain.

3. Is there a significant difference between the reading achievement of first grade students of different gender in relation to principal effects?

Null Hypothesis (H_{03}): There is no statistically significant difference in the growth trajectories of male and female students in relation to principal effects.

Research Hypothesis (H_{a3}): There is a statistically significant difference in the growth trajectories of male and female students in relation to principal effects.

3.1 There is no statistically significant difference in the growth trajectories of male and female students in relation to the PIQ Core Reading Program domain.

3.2 There is no statistically significant difference in the growth trajectories of male and female students in relation to the PIQ Leadership domain.

3.3 There is no statistically significant difference in the growth trajectories of male and female students in relation to the PIQ Professional Development domain.

3.4 There is no statistically significant difference in the growth trajectories of male and female students in relation to the PIQ Assessment domain.

3.5 There is no statistically significant difference in the growth trajectories of male and female students in relation to the PIQ Intervention domain.

4. Is there a significant difference between the reading achievement of first grade ethnic minority and non-minority students in relation to principal effects?

Null Hypothesis (H_{04}): There is no statistically significant difference in the growth trajectories of first grade ethnic minority and non-minority students in relation to principal effects.

Research Hypothesis (H_{a4}): There is a statistically significant difference in the growth trajectories of first grade ethnic minority and non-minority students in relation to principal effects.

4.1 There is no statistically significant difference in the growth trajectories of first

grade ethnic minority and non-minority students in relation to the PIQ Core Reading Program domain.

4.2 There is no statistically significant difference in the growth trajectories of first grade ethnic minority and non-minority students in relation to the PIQ Leadership domain.

4.3 There is no statistically significant difference in the growth trajectories of first grade ethnic minority and non-minority students in relation to the PIQ Professional Development domain.

4.4 There is no statistically significant difference in the growth trajectories of first grade ethnic minority and non-minority students in relation to the PIQ Assessment domain.

4.5 There is no statistically significant difference in the growth trajectories of first grade ethnic minority and non-minority students in relation to the PIQ Intervention domain.

5. Is there a significant difference between the reading achievement of first grade students with limited English proficiency and students that are fluent in English in relation to principal effects?

Null Hypothesis (H_{05}): There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to principal effects.

Research Hypothesis (H_{a5}): There is a statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to principal effects.

5.1 There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to the PIQ Core Reading Program domain.

5.2 There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to the PIQ Leadership domain.

5.3 There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in

relation to the PIQ Professional Development domain.

5.4 There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to the PIQ Assessment domain.

5.5 There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to the PIQ Intervention domain.

6. Is there a significant difference between the reading achievement of first grade students that qualify for exceptional student education and students that have no disability classification in relation to principal effects?

Null Hypothesis (H_{0_6}): There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to principal effects.

Research Hypothesis (H_{a_6}): There is a statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to principal effects.

6.1 There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to the PIQ Core Reading Program domain.

6.2 There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to the PIQ Leadership domain.

6.3 There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to the PIQ Professional Development domain.

6.4 There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to the PIQ Assessment domain.

6.5 There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to the PIQ Intervention domain.

Research Methods

Using necessary research elements suggested by Hallinger et al. (1996), this study included (a) a sufficient sample size, (b) a theoretically defensible model, (c) a sophisticated data collection instrument and (d) sophisticated data analysis tools. The use of these elements should capture the effects of principal behaviors on student achievement in reading (Hallinger et al., 1996; Witziers et al., 2003). In the following sections, a description of participant selection procedures is presented followed by the methods and data collection procedures.

Sampling Procedures

Principals of *Reading First* schools in Florida are required by the Florida Department of Education to complete an implementation survey in order to receive funding for the following year. Thus, the data set for this study was a criterion sample that included all of the 388 principals of *Reading First* schools in Florida in the 2004-2005 academic year.

Principal Population. Respondents on the PIQ were 388 elementary school principals in the state of Florida who were currently receiving *Reading First* funding. The age range for the sample was 21 to 67 ($M= 51.35 \pm 7.52$), with approximately 74% of the sample female ($n=285$). Sixty-seven (67) percent of respondents were Caucasian ($n=260$), 26% African American ($n=101$), 6% Hispanic ($n=22$) and <1% identified themselves as either Multi-racial ($n=3$) or Asian ($n=1$). The highest level of education completed for 72% of the principals was at the Master's level, followed by 15% Specialists, 12% Doctoral and 1% at the Bachelor's level. The average level of education experience for principals was 26.43 years ($SD=7.82$), with an associated average level of administrative experience at 12.49 years ($SD=6.81$). Principals reported a mean 5.77 years ($SD=4.95$) administrative experience at their current school and average 1.79 years ($SD=.41$) implementing the *Reading First* initiative at their current school. Finally, 90% of all participants completed the PIQ on an online format with the remaining 10% completing the instrument via paper and pencil.

The period of investigation for this study was from the first of four quarterly

Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Oral Reading Fluency (ORF) assessments for the 2004-2005 academic year, to the fourth ORF assessment of the 2004-2005 academic year. First grade students who received four continuous ORF assessments while residing in the same *Reading First* school were selected for this study.

Student Population. A total of 34,275 students from 367 elementary schools in the state of Florida were used in the present study. The sample was 52% male (n=17,882), 48% female (n=16,393), with all students currently enrolled in first grade. Thirty-six percent of students were African-American (n=12,350), 35% were Caucasian (n=12,035), 23% were Hispanic (n=7,848), 4% were Multi-racial (n=1,418), 1% were Asian (n=482) and <1% identified themselves as Native American (n=482).

Additional demographic statistics were tracked including (a) Socio-economic Status (SES), (b) Limited English Proficiency (LEP) and (c) Exceptional Student Education status (ESE). LEP statistics indicated that 80% of students did not have an applicable LEP code (n=27,447), 16% of students were LEP and enrolled in classes designed for LEP students (n=5,531), 3% were students being followed for a two year period after having exited from the ESOL program and <1% of students were either LEP and not enrolled in specifically designed classes (n=8), students for whom a two-year follow up period had been completed after the student had exited the ESOL program (n=81), or data was missing (n=68).

SES results indicated that 65% of students were eligible for free lunch (n=22,371), 20% did not apply for free or reduced price lunch (n=6,791), 10% were eligible for reduced price lunch (n=3,457), 4% were not eligible for either free or reduced price lunch (n=1,476) and < 1% were either enrolled in a USDA-approved Provision Z school (n=116) or had missing data (n=64).

ESE results indicated that 85% of participants did not have an applicable exceptionality status (n=29,056). Six percent of students were individuals with a speech impairment (n=2,188), 3% were individuals with a language impairment (n=1,026), 3% were individuals with a specific learning disability (n=965) and < 1% of the sample were students with a variety of disabilities (n=1,040).

Data Selection

This study examined the relationship between the principal's role in

implementation of elements of effective reading programs and student achievement in reading. Data selected to analyze this relationship included (a) principal's responses to the 2004-2005 *Principal Implementation Questionnaire* (PIQ) and (b) the results of first grade ORF assessments for all selected first grade students. Longitudinal student comprehension data were not available for this population, so data derived from four consecutive measures of oral reading fluency were used to determine levels of reading achievement. The rationale supporting this choice stems from research that consistently identifies student scores on assessments of oral reading fluency as strong predictors of overall reading comprehension, which is the end goal of reading instruction (Fuchs, Fuchs, Hosp & Jenkins, 2001; Good, Simmons & Kame'enui, 2001). Indeed, comprehension has come to be recognized as the essence of reading, essential for academic learning in all subject areas and in overall lifelong learning (Durkin, 1993; National Reading Panel, 2000).

Measures and Reliability

The effects of principals on school climate, school effectiveness and student achievement have been of interest for several decades (Ediger, 1999; Hallinger & Heck, 1996; Levine & Lezotte, 1990; Witziers, Bosker & Kruger, 2003). Furthermore, the role of the principal as the curriculum leader, a guide, a model and a mentor (Ediger, 2000) has been addressed by many scholars in the field (e.g., Sammons, Hillman & Mortimore, 1995; Levine & Lezotte, 1990). The principal is one of the key factors that determine the success of various interventions in a school. Particularly, the success of a reading program heavily depends on principal leadership.

While it is clear that principals influence overall school effectiveness, the specific behaviors associated with student achievement are unclear (Hallinger & Heck, 1996, 1998; Pitner, 1988; Snyder, Acker-Hocevar & Snyder, 1996; Witziers et al., 2003). Thus, the effects of instructional leadership on individual student learning remain a complex and unanswered question (Hallinger & Heck, 1996, 1998; Heck, 1993; Bossert, Dwyer, Rowan & Lee, 1982). Moreover, since reading is the primary skill affecting student development in all cognitive and social domains (Snow, Barnes, Chandler, Goodman & Hemphill, 1991; Snow, Burns & Griffin, 1998), principals are expected to be

knowledgeable in methods of reading instruction.

Researchers have yet to agree upon appropriate methods of assessing leadership behaviors, as leadership is a multifaceted and contextually-based construct (Hallinger et al., 1996; Heck & Larson, 1996; Witziers, Bosker & Kruger, 2003). A lack of consensus in operationalizing leadership has impeded efforts to adequately capture the effects of leadership on student achievement (Bossert, 1988; Hoy & Ferguson, 1985; Witziers et al., 2003). As comprehensive assessment of principal leadership practices is daunting—especially in relation to student achievement—it is all the more imperative that valid and reliable measures be developed. The PIQ was constructed to measure the quality of principal implementation practices in relation to effective reading programs.

The PIQ is a 105-item Web-based inventory that was given to school principals to inquire about the implementation of *Reading First* activities in their school during the 2004-2005 school year. There are five dimensions of the PIQ that should be completed in approximately 25 minutes by school principals: (a) Core Reading Program, (b) Professional Development, (c) Leadership/Organizational Practices, (d) Assessment and (e) Intervention. These five dimensions were selected and developed based on priorities set forth by state and federal *Reading First* policies, as well as a research-based understanding of the school-level elements needed to promote student achievement in reading. Further, each of these dimensions is focused on elements of overall reading program implementation that can be influenced by principal behaviors. Operational definitions for these five dimensions are presented below.

Core Reading Program

Implementation of a systematic core reading program is essential to a successful reading program. Current core reading programs are designed using scientifically-based reading research as a foundation for instructional practice. The PIQ includes questions regarding the quality of core reading program implementation in each school that can be influenced by principal behavior.

Professional Development

Providing effective professional development that is targeted to the actual needs of school staff is one of the responsibilities of any principal and, particularly, *Reading First* principals. The challenges faced by *Reading First* schools are magnified due to

especially challenging student populations, adding emphasis to the need for effective professional development. Principals should provide opportunities for professional development to build the professional competency of staff and themselves, improve environmental conditions effecting reading instruction and build rapport and collegiality among faculty. Thus, the PIQ includes questions regarding the existence and quality of professional development efforts in each school.

Leadership/Organizational Practices (Leadership)

Clearly, effective leadership is critical to the implementation of effective reading programs at the school level, as it is in any school improvement effort. Organizational management, goal setting, progress monitoring and evaluation are some of the essential elements associated with this dimension. In this section of the PIQ, principals were asked to respond to items that are focused on the relevant practices that are associated with effective instructional leadership in reading.

Assessment

Developing an appropriate assessment plan, monitoring student outcome data and using these data to improve overall instructional practices are elements of this dimension. The use of assessment to inform and drive reading instruction is a keystone of *Reading First* and, indeed, any instructional effort. This section of the PIQ includes items associated with the principal's role in ensuring a powerful assessment component is part of overall program implementation.

Intervention

One of the most compelling aspects of *Reading First* is the requirement that schools implement an Immediate Intensive Intervention program based on credible reading research. These programs are designed to provide focused, explicit reading instruction to struggling readers on their specific instructional level. Principal oversight of this aspect of reading implementation is crucial because effective intervention is, perhaps, the most powerful tool in the overall improvement of school-level reading achievement, while being the least understood from a practical standpoint. Thus, the PIQ includes items associated with the principal's role in intervention.

Validation Process

The PIQ was distributed to all principals of *Reading First* schools in Florida by

the Florida Center for Reading Research (FCRR) at Florida State University via the Progress Monitoring and Reporting Network (PMRN) maintained by the center. The Florida Center for Reading Research is part of the Florida *Reading First* Leadership Triangle and is the designated evaluator of statewide *Reading First* programs. The PMRN is a data collection and reporting system through which *Reading First* schools both report reading data and receive reports on these data. Principals, reading coaches and teachers in *Reading First* schools use this system as the primary source of *Reading First* data regarding the state, districts, schools, classrooms and students.

Principals were prompted via email to log on to the PMRN and complete the survey instrument. The PIQ Web-page on the PMRN was active for a period of one week. Prior to activation, principals were informed via email of the specific dates of the response period. Data derived from the survey were automatically compiled into a database formatted for empirical analysis.

The PIQ employs primarily Likert-based items, with some qualitative free-response items in the last section of the survey. The PIQ is a revised version of the 2003-2004 Principal Implementation Survey that was previously validated using appropriate psychometric methods. From these analyses, the survey was revised to improve the reliability and validity of the instrument. For purposes of this study, internal consistency analyses and confirmatory factor analyses were conducted to determine reliability and validity of the individual items within the PIQ. Therefore, only those items determined to be valid and reliable were used for the final analyses. The paper and pencil version of the PIQ is provided as Appendix A.

For the validation phase of this study, responses were received from 352 elementary school principals in the state of Florida who were currently receiving *Reading First* funding. The complete sample of 388 was not used because the validation analysis was initiated before all of the surveys were received, due to external time limitations. The age range for the sample was 21 to 67 ($M= 51.48 \pm 7.46$), with approximately 73% of the sample female ($n=257$). Participants were 68% Caucasian, 25% African American, 6% Hispanic and 1% identified themselves as Multi-racial. The highest level of education completed for 74% of the principals was at the Master's level, followed by 13% Specialist's, 12% Doctoral and 1% at the Bachelor's level. The average level of

education experience for principals was 26.51 years (SD=7.73), with an associated average level of administrative experience at 12.58 years (SD=6.86). Principals reported a mean 5.77 years (SD=5.01) administrative experience at their current school and average 1.79 years (SD=.41) implementing the *Reading First* initiative at their current school.

Table 1 provides a summary of PIQ domains that proved to be reliable measures. As displayed, four of the original five dimensions were retained for the empirical study (Professional Development, Leadership, Assessment and Intervention). A total of 28 items were retained across the four dimensions. Combined, the four dimensions display an overall alpha reliability of .92.

Table 1.
Descriptive Statistics for PIQ Results

Factor ^a	Mean	Standard Deviation	SEM ^b	α
PROFD (n=2)	7.17	1.12	.48	.81
LEAD (n=8)	29.28	3.36	1.11	.89
ASSESS (n=13)	45.10	5.59	1.77	.90
INTERV (n=5)	14.79	3.40	1.63	.77
Total (n=28)	96.31	10.38	2.94	.92

^a PROFD = Professional Development, LEAD = Leadership, ASSESS = Assessment, INTERV = Intervention. ^b SEM= Standard error of measure

Descriptive measures of overall model fit for the confirmatory factor analysis (CFA) performed in LISREL 8.7 are provided in Table 2. As is shown, five iterations of model configuration were conducted to identify the best fitting model from which to conduct the CFA, with the fifth attempt resulting in the model used for this study.

The first group of statistics in Table 2 compare individual components of the selected data to the overall data population. The χ^2/df statistic is traditionally considered to be a benchmark measure when considering appropriate model fit. In this case, the

selected model displayed an appropriate result for this measure (3.0). However, due to the sensitivity of the χ^2/df statistic to sample size, other goodness-of-fit measures have been developed to provide a more comprehensive understanding of model fit. Collectively, these measures of overall model fit indicate how well a structural equation model corresponds to the data being analyzed and, thus, how appropriate the results are in drawing conclusions from the data.

The Root Mean Square Error of Approximation (RMSEA; Steiger, 1990) is also a measure of approximate fit in the population. Steiger (1990) defines a close fit as a RMSEA value less than or equal to .05. Browne and Cudeck (1993) suggest that RMSEA values $\leq .05$ can be considered as a good fit, values between .05 and .08 as an adequate fit and values between .08 and .10 as an acceptable fit. Values greater than .10 are not considered acceptable. The selected model has a value of .08 and is, therefore, considered acceptable.

The Root Mean Square Residual Index (RMR) developed by Jöreskog and Sörbom (1981, 1989) is another widely accepted measure of model fit. RMR values approaching zero suggest a good fit, as a value of zero indicates a perfect fit. In general, the RMR should be less than .05 for a good fit (Hu & Bentler, 1995), whereas values smaller than .10 can be interpreted as acceptable. The value of .03 displayed by the selected model indicates an especially good fitting model, based on this statistic (see table 2).

The next group of statistics provided in Table 2 is comprised of descriptive measures based on model comparisons. The basic idea of such comparison indices is that the fit of a selected model is compared to the fit of some baseline model to provide a clearer indication of overall appropriateness. The Normed Fit Index (NFI; Bentler & Bonnett, 1980) includes values that range from zero to one. Higher values indicate a better fitting model in comparison to a baseline model. For this index, .95 or higher indicates good fit, while values greater than .90 should be interpreted as indicating an acceptable fit (Schumacker & Lomax, 1996). The .92 value shown in Table 2 is indicative of an acceptable fitting model.

Table 2.
Summary of Model Fit Indices for CFA Iterations

Model	Items	χ^2/df	RMSEA	90% CI ^a	RMR	NFI	CFI	GFI	AGFI
1	58	4.6	.102	(.099, .104)	.07	.86	.90	.58	.55
2	43	6.0	.120	(.121, .123)	.05	.88	.90	.59	.55
3	38	3.5	.085	(.082, .089)	.05	.91	.94	.74	.71
4	33	3.2	.084	(.080, .089)	.03	.91	.92	.79	.76
5*	28	3.0	.080	(.073, .082)	.03	.92	.94	.81	.77

Note: RMSEA= root mean square error of approximation; RMR= root mean residual; NFI= normed fit index; CFI= comparative fit index; GFI= goodness of fit index; AGFI= adjusted goodness of fit index.

* Final model retained

^a Represents the 90% confidence interval for RMSEA

The Comparative Fit Index (CFI; Bentler, 1990) also ranges from zero to one with higher values indicating better fit. This index is typically interpreted as a value of .97 being indicative of good model fit, while values greater than .95 generally being interpreted as an acceptable fitting model. The GFI is one of the most stable fit indices because it has been shown to be less affected by sample size (Bentler, 1990; Bollen, 1990; Hu & Bentler, 1995). While the .94 value associated with the selected model is just below the acceptable fit range in this case, this statistic should be taken into the context of other more favorable values associated with this model.

The Goodness-of-Fit-Index (GFI; Jöreskog & Sörbom, 1989) typically ranges between zero and one with higher values indicating better fit. However, there are some cases that may result in a negative GFI. The usual rule of thumb for this index is that .95 is indicative of good fit relative to the baseline model, while values greater than .90 are usually interpreted as indicating an acceptable fit (Schumacker & Lomax, 1996). The

value of .81 is, again, below acceptable range, but should be considered in the overall favorable statistical context.

Jöreskog and Sörbom (1989) also developed the Adjusted Goodness-of-Fit Index (AGFI) with values that typically range between zero and one. Larger values indicate a better fit, but it is also possible to have a negative AGFI in some situations. Generally, .90 indicates a good fit, while values greater than .85 should be considered an acceptable fit. In this model, the AGFI is .77. This lower value is considered reasonable in light of the acceptable values associated with the previous fit indices for the selected model.

Table 3 provides detail in the structure coefficients for the four remaining PIQ constructs. Coefficients for each of the PIQ items retained for the analysis are included.

A correlation matrix that compares the PIQ constructs to each other is provided in Table 4. As is shown, the Assessment construct is significantly correlated to Leadership, while the Intervention construct is significantly correlated to all other constructs.

The Dynamic Indicators of Basic Early Literacy Skills (DIBELS; University of Oregon, 2005) are a set of standardized, individually administered measures of early literacy development. The instruments are designed to be short (approximately one minute) fluency measures used to regularly monitor the development of pre-reading and early reading skills. The DIBELS measures were developed at the University of Oregon using the essential domains of early reading discussed in both the National Reading Panel (2000) and National Research Council (1998) reports to assess student development of (a) phonological awareness, (b) alphabetic understanding and (c) automaticity and fluency with text. Each measure has been thoroughly researched and demonstrated to be reliable and valid indicators of early literacy development (Good & Kaminski, 2002; Good, Simmons & Kame'enui, 2001; University of Oregon, 2005). Further, the measures have been demonstrated to be predictive of later reading proficiency to aid in the early identification of students who are progressing below expectations. Thus, the DIBELS Oral Reading Fluency (ORF) data collected for this study, using the methods prescribed by the instrument publisher, are assumed to be valid and reliable indicators of student achievement in oral reading fluency.

Table 3.
 Estimated Structure Coefficients for PIQ Constructs

Item	Professional				R ²
	Leadership	Development	Assessment	Intervention	
13		.97			.95
14		.71			.50
16	.74				.55
17	.72				.52
18	.76				.57
19	.74				.54
22	.71				.51
23	.68				.46
24	.66				.44
29	.66				.43
30			.52		.27
31			.52		.27
32			.50		.25
57			.59		.35
64			.66		.44
65			.56		.32
66			.78		.61

Table 3 , cont.

Item	Professional			R ²
	Leadership	Development	Assessment	
67			.74	.55
68			.75	.56
69			.59	.35
70			.71	.51
71			.71	.51
72			.71	.50
80				.40
81				.47
82				.49
83				.37
86				.26

^a Estimates for constructs are statistically significant at $z > 2$ level.

The ORF subtest is a classroom-level, standardized, individually administered test of accuracy and fluency with connected text. The ORF passages and procedures are based on the program of research and development of Curriculum-Based Measurement of Reading (CBM) by the University of Minnesota (Shinn, 1989). The measure is comprised of a standardized set of passages and administration procedures designed to (a) identify students who may need additional instruction and (b) monitor progress toward instructional goals (Fuchs, Fuchs, Hosp & Jenkins, 2001). The ORF passages are calibrated for the appropriate goal level of reading for each grade in which it is

Table 4.
Correlations for PIQ Dimensions

Factor ^a	LEAD	PROFD	ASSESS	INTERV
LEAD	1.00			
PROFD	.001	1.00		
ASSESS	.625*	.040	1.00	
INTERV	.311*	.151*	.480*	1.00

^a LEAD = Leadership, PROFD = Professional Development, ASSESS = Assessment, INTERV = Intervention.

* Correlation statistically significant at $p < .05$ level.

administered (grades 1-3). Student performance in oral reading fluency is measured by having students read a passage aloud for one minute. From this running reading record, words omitted, substituted and hesitations of more than three seconds are scored as errors. Words self-corrected within three seconds are scored as accurate. The number of correct words per minute from the passage is considered the oral reading fluency rate (Good, Simmons & Kame'enui, 2001; University of Oregon, 2005).

Florida student oral reading fluency data are housed in a database maintained by FCRR and was available for use in this study. All required protocol for accessing ORF data via the PMRN were followed in the course of this study.

Data Summary and Analysis Procedures

Data are summarized and reported according to (a) the scores on the four remaining sub-scale dimensions of the PIQ and (b) the four instances of the ORF subtest collected from all *Reading First* schools in Florida during the 2004-2005 academic year. The primary methodology that was used to investigate principal-level factors affecting student achievement in reading and growth in student reading achievement was

Hierarchical Linear Modeling (HLM); (Raudenbush & Bryk, 2002; Raudenbush, Bryk, Cheong & Congdon, 2001). HLM is widely utilized for separating student and school effects on student achievement and is often used to identify the discrete impact of programs at the school level. In fact, HLM appears to be generally replacing multiple linear regression and repeated measures techniques as the method of choice for school-effects research (Mendro, Webster, Bembry & Orsak, 1995; Meyer, 1997).

Raudenbush and Bryk (2002) developed HLM and cite three primary purposes for the approach:

...improved estimation of effects within individual units (e.g., developing an improved estimate of a regression model for an individual school by borrowing strength from the fact that similar estimates exist for other schools); the formulation and testing of hypotheses about cross-level effects (e.g., how varying school size might affect the relationship between social class and academic achievement within schools); and the partitioning of variance and covariance components among levels (e.g., decomposing the covariation among sets of student-level variables into within- and between-school components; p. 7).

Hierarchical Linear Modeling (HLM) allows researchers to overcome many of the traditional difficulties associated with less complex statistics analyses. Single-level analyses, as are typically incorporated, cannot adequately partition variance to appropriate groups and prevent the separation of individual and group effects on the outcome of interest, in this case student reading achievement. On the other hand, HLM is designed to separate individual and group effects on the outcome of interest, without the traditional difficulties of single-level analyses.

Longitudinal student achievement data were analyzed for principal effects utilizing growth curve modeling, a procedure that has been used by researchers in numerous fields for the study of intra-individual differences in change over time (Bryk & Raudenbush, 1992; Rogosa, Brandt & Zimowski, 1982; Willett & Sayer, 1994). Growth curve models can be viewed as falling within the class of multilevel linear models (Bryk & Raudenbush, 1992). Thus, this study used the Hierarchical Linear Modeling (HLM 6) statistical software package for all modeling analyses.

A three-level growth curve model was tested to determine the amount of student-

level variance that can be explained by the four dimensions measured by the PIQ. Level 1 of this model tested the growth of ORF scores over time for each of the four assessments at the individual level. Level 2 modeled selected student-level variables (i.e., minority status, socio-economic status as determined by free and reduced lunch eligibility, whether students had limited English proficiency and whether a student was classified as disabled). Level 3 modeled variables associated with the principal. Principals' responses on the five dimensions of the PIQ were used to define and categorize principal behavior in relation to the implementation of effective reading programs. Figure 1 provides a graphical representation of the variable relationships modeled in this study.

The three levels were modeled to determine the amount of student-level variance that could be attributed to time, intra-individual characteristics and principals' efforts to implement effective reading programs in their schools, as measured by the PIQ. This was a direct effect model that attributed an appropriate portion of the total student-level variance in ORF performance to the student- and principal-level variables. In addition, the analysis identified the proportion of unexplained residual variance that the model could not account for.

Another important aspect of this analysis was that it modeled the principal effects on the reading achievement of students over time. The longitudinal data set provided opportunities to measure principal effects on reading growth across four separate, consecutive assessments that encompassed a full academic year. From this analysis, student growth trajectories were formulated and the impact of principal implementation of effective reading programs was analyzed over time to identify changes in growth.

Figure 1 illustrates the model of analysis utilized for this study. Principal responses to the four dimensions of the PIQ (LEAD, PROFD, ASSESS and INTERV) comprise the latent PIQ construct. This construct was then used to model the intercepts and slopes associated with the four instances of ORF testing (ORF1, ORF2, ORF3 and ORF 4).

All demographics and calculated results are reported using appropriate statistical tables to provide for a comprehensive review of the data. In addition, growth curve trajectories are provided in chart format to illustrate the growth of student achievement in reading as related to the identified principal implementation behaviors.

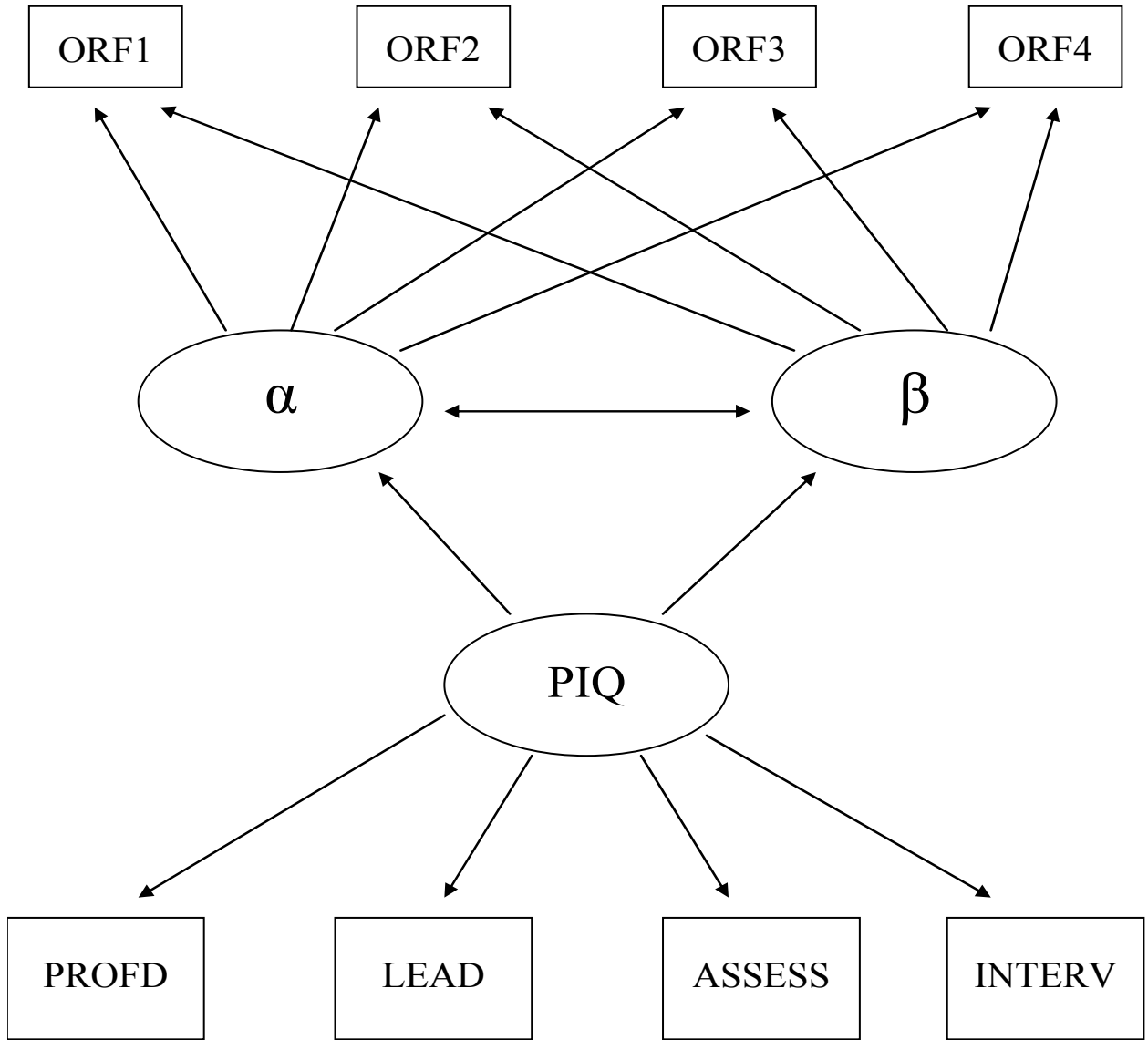


Figure 1.
 Analysis Diagram for Three-level Growth Curve Model

CHAPTER 4

FINDINGS

All of the data collected via the methods and procedures as specified in Chapter Three are presented in this chapter. The data are presented in two categories: (a) research data and (b) research questions and hypotheses. The research data section presents descriptive results from the analytical components of the study that built the necessary framework from which to construct the Level 3 hierarchical model, which models the principal effects on student reading achievement as measured by the DIBELS assessment of oral reading fluency (ORF). The data included (a) responses to the Principal Implementation Questionnaire (PIQ), (b) student reading performance (ORF) by subgroup and (c) statistical results of the Analysis of Variance (ANOVA), Level 1 and Level 2 models used to provide the necessary rationale to support Level 3 modeling. Research questions and hypotheses are then addressed utilizing results from the Level 3 model.

Research Data

PIQ Responses

The Principal Implementation Questionnaire (PIQ) initially included five dimensions that assessed the quality of principal implementation of effective reading programs. Four were retained for analysis (Professional Development, Leadership, Assessment and Intervention), based on statistical validation explained in Chapter Three. A total of 28 survey items were retained—across the four dimensions—on a four-point Likert-type scale, ranging from R1 (*strongly disagree* or *never*, depending on the item) to R4 (*strongly agree* or *always*, depending on the item). The result was a mean score for each respondent on each dimension of the PIQ. The total possible scores for each dimension were (a) Professional Development (8), (b) Leadership (40), (c) Assessment (52) and (d) Intervention (20). A higher value response on PIQ items, and thus, a PIQ

domain, is considered as a more favorable level of perceived reading program implementation by principals. Table 5 displays the retained items comprising each of the PIQ dimensions as well as the mean principal response and the percentage frequency responses for each item. The survey item numbering in Table 5 represents the actual item number within the PIQ and begins at item 13, ends at item 86 and omits many sequential numbers in between. While these PIQ items are shown as configured in the previous instrument, future PIQ assessments will use the validated format, containing only the retained items.

The mean response selected by principals (n = 388) for each of the 28 items was 3.25. The frequency responses across each of the 28 PIQ items were dispersed as follows: (a) 2.88% of principals selected R1, (b) 9.56% of principals selected R2, (c) 31.80% of principals selected R3, and (d) 56.86% of principals selected R4, indicating that a majority of responses were “strongly agree” or “always” in nature. While these results do not address the research questions of this study in isolation, they are the principal-level data that are used in comparison to student achievement data in the Level 3 model. They are also important to the overall understanding of principals’ responses on the PIQ regarding their implementation of effective reading programs.

Population Subgroup ORF Data

First grade students in Reading First schools were assessed in oral reading fluency using the corresponding DIBELS subtest (ORF). This oral assessment measures correctly pronounced words per minute. Oral reading fluency was selected for this study because it has been shown to be directly related to reading comprehension, the overall goal of reading instruction. The following tables (Tables 6-10) provide ORF scores across the four instances of assessment disaggregated by the selected subgroups including (a) socioeconomic status (SES), (b) gender, (c) ethnicity, (d) limited English proficiency status (LEP) and (e) students qualifying for exceptional student education (ESE). Again, these results alone do not address the central research questions, but are critical in addressing these questions as part of the Level 3 model.

Table 5.

Descriptive Statistics for Retained PIQ Items by Construct

Survey Item	<i>M</i>	<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>R4</i>
PROFESSIONAL DEVELOPMENT (n = 2)					
13. Professional development for teachers in implementing effective reading programs has been essential to improving student achievement in reading.	3.67	1.5	0.5	27.1	70.6
14. Professional development for principals in implementing effective reading programs has been essential to improving student achievement in reading.	3.50	1.5	3.4	38.4	56.4
LEADERSHIP (n = 8)					
16. I provided opportunities for teachers to use reading assessment data to guide instructional decisions.	3.73	0.8	-	25.1	74.2
17. I monitored teachers' use of assessment data to guide reading instruction.	3.59	1.3	1.0	35.3	62.1
18. I was successful in finding individuals in our school to help classroom teachers in providing interventions to our most struggling readers in K-3.	3.63	1.0	2.3	29.4	67.0
19. We have successful intervention programs in place in our school for our most struggling readers in grades K-3.	3.54	1.0	3.9	35.3	59.5
22. The school schedule was constructed to provide sufficient time for reading intervention.	3.77	0.8	1.0	18.6	79.4
23. The school schedule was constructed to provide time for staff members other than classroom teachers to deliver reading interventions.	3.60	1.0	4.9	27.1	66.9
24. The school schedule was constructed to provide time for teachers to collaborate on reading issues.	3.59	1.0	3.6	30.2	64.9
29. I actively promoted the full implementation of reading programs in my school.	3.79	0.3	-	20.1	79.6
ASSESSMENT (n = 13)					
30. Our assessment team was essential to overall school-wide reading performance this year.	3.61	0.5	2.8	31.7	64.9
31. We currently have an effective plan for implementation of <i>Reading First</i> Assessment.	3.62	0.5	1.8	32.7	64.9
32. A comprehensive screening process is currently in place to assess student progress in reading.	3.64	0.3	2.8	29.1	67.8
57. I provided opportunities for teachers to give feedback on the implementation of effective reading programs in my school.	3.00	-	7.5	49.7	42.5

Table 5, cont.

Survey Item	<i>M</i>	<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>R4</i>
64. School assessment data were used to allocate appropriate resources for critical instructional areas.	3.41	1.5	9.0	36.3	52.8
65. Stakeholder feedback was monitored to assess the effectiveness of the school assessment plan.	2.94	3.4	26.3	42.8	27.3
66. We used assessment data to provide feedback to teachers regarding instructional practices.	3.55	-	5.7	33.8	60.3
67. We used assessment data to determine professional development needs of staff.	3.54	-	5.7	34.5	59.5
68. We used assessment data to determine which intensive instructional practices were most effective.	3.38	0.8	11.6	36.6	50.8
69. We used assessment data to determine needed support from outside sources (i.e., district office, FCRR, Just Read, Florida!).	3.01	5.9	22.9	35.1	35.8
70. We used assessment data to determine growth patterns in struggling students' reading achievement.	3.57	-	4.6	34.0	61.1
71. We used assessment data to determine students' strengths and weaknesses in reading.	3.71	-	1.8	25.3	72.7
72. We used assessment data to plan for implementation of interventions for struggling readers.	3.71	-	2.1	25.3	72.4
INTERVENTION (n = 5)					
80. We used computer software to provide interventions <i>outside</i> of the regularly scheduled reading block.	3.10	3.9	22.4	33.5	39.9
81. We used ESE teachers to provide additional instruction in reading <i>outside</i> of the regularly scheduled reading time.	2.89	10.3	24.3	31.3	34.1
82. We used resource teachers to provide additional instruction in reading <i>outside</i> of the regularly scheduled reading time.	2.93	10.9	21.2	31.8	36.2
83. We used paraprofessionals to provide additional instruction in reading <i>outside</i> of the regularly scheduled reading time.	2.76	11.1	31.7	27.3	29.6
86. We used published intervention programs to help guide interventions for our struggling students.	3.07	4.1	23.7	33.0	38.9

Note: Estimates in R1-R4 are percentage frequency responses.

These data are reported to illustrate the similarities among student subgroups that comprise Level 2 of the growth curve model in relation to growth in oral reading fluency.

Socio-economic Status (SES)

In general, ORF scores increased for all SES categories across the four instances of ORF assessment. There were minimal differences between students that were eligible for free/reduced lunch (n=22,371) and those that were not eligible. For example, the ORF1 means ranged from a low of 16.58 (EFL) to a high of 26.54 (EZS). Students with free and/or reduced lunch status were consistently the lowest two categories with regard to ORF growth across all four instances of ORF assessment (see Table 6).

Table 6.
Means and Standard Deviations of Socio-economic Status as Indicated by Free/Reduced Lunch Eligibility

Classification	ORF1		ORF2		ORF3		ORF4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ANE	23.76	21.85	35.42	27.58	46.06	33.15	54.57	32.62
EFL	16.58	15.93	26.23	21.94	33.83	27.10	43.16	28.19
ERL	20.63	19.13	31.93	24.85	40.90	30.45	50.28	30.63
EZS	26.54	22.95	39.90	28.28	49.15	33.31	56.04	31.90
DNA	24.93	22.64	37.88	28.42	48.68	33.95	57.91	33.63

Note: ANE = student applied for free or reduced lunch but not eligible, EFL = student is eligible for free lunch, ERL = student is eligible for reduced lunch, EZS = student is enrolled in provision Z school, DNA = student did not apply for free or reduced lunch.

Gender

Oral reading fluency scores are presented by gender in Table 7. As can be seen, female students in the study population outperformed males in each of the four ORF assessments during the 2004-2005 school year. Male students were shown to be consistently lower in ORF mean scores than female students across each of the four ORF

assessments. The difference between the mean scores at ORF1 was 4.17 words per minute (WPM); growing to 7.99 WPM at ORF4 (see Table 7). At ORF1, males students displayed a mean WPM of 16.99; however, the standard deviation was a higher value ($SD=17.08$). This reflects a platykurtic distribution, that is, a distribution that has thin tails and a relatively flat middle. Compared with a normal distribution, a larger proportion of its observations are clustered within two standard deviations of the mean. In this case, the platykurtic distribution means that there was less variation in ORF scores for male students at ORF1, violating the assumption of a normal distribution of ORF scores.

Table 7.
Means and Standard Deviations by Gender

Classification	ORF1		ORF2		ORF3		ORF4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Males	16.99	17.08	26.42	22.56	34.41	27.97	43.51	29.10
Females	21.16	19.54	32.98	25.85	42.05	31.31	51.50	31.26

Ethnicity

Within the population of interest, 34,275 students fell into ethnic minority categories. In general, the Asian-American students in the population completed the most words correctly per minute, followed by Multiracial students, Caucasian students, Native American students, African American students and Hispanic students (see Table 8). Among the ethnic categories, growth in ORF was similar and relatively stable. However, the difference in mean scores between the lowest category (Hispanic) and the highest (Asian Americans) was 10.07 WPM at ORF1, but decreased to 9.39 at ORF4. This shows a slight closing of the gap between high and low students in the Ethnicity subgroup across the assessment periods (Table 8).

Table 8.
Means and Standard Deviations by Ethnicity

Classification	ORF1		ORF2		ORF3		ORF4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Caucasian	21.58	20.62	33.11	26.65	42.66	32.17	51.93	32.45
African American	17.85	16.76	28.04	22.68	35.68	27.84	44.53	28.32
Asian American	25.90	23.19	39.05	29.09	51.74	35.20	61.97	34.80
Hispanic	15.83	15.37	25.17	21.32	33.07	26.89	42.87	28.30
Native American	18.79	18.06	30.96	24.99	39.13	28.98	49.75	30.68
Multiracial	21.84	21.68	33.46	27.39	42.72	32.50	52.13	32.30

Limited English Proficiency (LEP)

The opposite effect was seen among the categories of the LEP subgroup (n = 6,828). As shown in Table 9, LEP students not enrolled in LEP classes (LN) and students being followed up for a two year period after exiting an ESOL program (LF) displayed the highest ORF scores, followed by students whose ESOL follow up was completed (LZ) and LEP students enrolled in LEP classes (LY). The difference between the lowest category at ORF1, LEP students in LEP classes (LY; 13.04 WPM) and the highest category, LEP students not enrolled in LEP classes (LN; 22.88 WPM), was 9.84 WPM at ORF1. The mean score for all categories was 18.99. This gap widened to 12.29 WPM at ORF4 between the lowest students (LY; 38.22 WPM) and the highest students, those who were being followed up for two years after exiting ESOL programs (LF; 50.51 WPM). The mean score for all categories at ORF4 was 45.61 (Table 9). Once again, the table identifies three platykurtic distributions, meaning there was little variation in ORF scores in these circumstances.

Table 9.
Means and Standard Deviations by LEP Classification

Classification	ORF1		ORF2		ORF3		ORF4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
LY	13.04	13.19	21.35	18.96	28.48	24.15	38.22	26.20
LN	22.88	30.03	30.63	30.72	41.63	40.32	49.13	39.89
LF	20.14	16.44	31.19	22.39	40.29	28.27	50.51	29.18
LZ	18.78	14.41	28.30	20.49	33.83	24.60	41.12	22.67
ZZ	20.14	19.63	31.16	25.14	39.93	30.60	49.07	30.94

Note: LY = Student is LEP and enrolled in classes designed for LEP students, LN = student is LEP and not enrolled in classes designed for LEP students, LF = student is being followed up for two year period after exiting ESOL program, LZ = student's ESOL follow up has been completed, ZZ = not applicable.

Students Qualifying for Exceptional Student Education (ESE)

ORF scores by ESE classification (n=5,219) are provided in Table 10. Six percent (6%) of the total student population were individuals with a speech impairment (n=2,188), 3% were individuals with a language impairment (n=1,026), 3% were individuals with a specific learning disability (n=965) and < 1% of the sample were students with a variety of disabilities (n=1,040). That being said, only the three categories comprised of a significant proportion of students (SI, LI and SLD) were considered for discussion.

In general, ORF scores of each category of the ESE subgroup increased across the four instances of ORF assessment. The difference between the lowest category at ORF1, students with specific learning disabilities (SLD; 10.16 WPM) and the highest category, students with speech impairments (SI; 16.45 WPM), was 6.29 WPM at ORF1. The mean score for all categories was 13.38. This gap widened to 16.77 WPM at ORF4 between the lowest students (SLD; 25.49 WPM) and the highest students, (SI; 42.26 WPM). The mean score for all categories at ORF4 was 34.10 (Table 10). Platykurtic distributions are

shown for many of the ESE classification categories.

Table 10.
Means and Standard Deviations by ESE Classification

Classification	ORF1		ORF2		ORF3		ORF4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
EMH	4.22	9.54	6.18	12.31	7.76	13.52	11.56	17.01
TMH	0.00	0.00	0.08	0.29	0.50	1.45	1.75	3.98
OI	15.09	19.78	23.82	25.93	28.69	31.27	36.02	35.21
SI	16.45	15.67	25.39	21.38	32.95	26.39	42.26	27.56
LI	13.54	13.63	20.55	19.16	26.43	23.95	34.54	26.18
DHH	10.36	7.42	15.43	14.73	19.00	18.43	24.39	22.42
VI	22.27	26.93	28.45	40.40	36.27	43.94	48.18	49.76
EH	12.20	14.43	17.78	20.64	22.55	25.41	28.41	27.54
SLD	10.16	18.85	14.96	15.62	18.89	18.81	25.49	21.09
GIF	63.21	34.61	81.77	36.06	97.58	36.73	101.03	35.21
HH	11.80	7.53	18.60	7.62	26.40	7.60	38.60	11.52
PMH	4.00	-	9.00	-	19.00	-	23.00	-
AUT	27.58	38.10	32.26	35.23	40.32	43.29	47.87	41.66

Table 10, cont.

Classification	ORF1		ORF2		ORF3		ORF4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SED	8.75	7.29	14.25	11.61	15.95	12.81	26.70	18.71
TBI	16.29	25.87	28.29	28.07	37.86	41.46	50.71	43.07
DD	13.39	16.02	21.68	20.42	28.88	26.28	38.29	28.65
EC	14.00	-	19.00	-	24.00	-	50.00	-
OHI	11.89	15.87	17.49	19.25	22.28	25.86	30.21	29.45
NA	19.64	18.37	30.78	24.34	39.69	29.78	49.20	30.09

Note: EMH = educable mentally handicapped, TMH = trainable mentally handicapped, OI = orthopedically impaired, SI = speech impaired, LI = language impaired, DHH = deaf or hard of hearing, VI = visually impaired, EH = emotionally handicapped, SLD = specific learning disability, GIF = gifted, HH = hospital/homebound, PMH = profoundly mentally handicapped, AUT = autistic, SED = severely emotionally disturbed, TBI = traumatic brain injured, DD = developmentally delayed, EC = established conditions, OHI = other health impaired, NA = not applicable

Three-Level Growth Curve Model

Model Formulation

There have been many advances in the use of multilevel data to model educational outcomes. The use of HLM to investigate the interactions of endogenous and exogenous variables on student outcomes is increasingly becoming the methodology of choice because of its ability to simultaneously analyze data nested in naturally occurring groups. Further, since students are not randomly assigned to schools and teachers, multilevel methods provide the most precise and rigorous analyses of interactions affecting student performance because they address student-level and school-level variables in naturally-occurring situations (Phillips & Adcock, 1997). As a result, decisions in educational policy are frequently based on the findings of HLM analyses.

Within a typical school structure, students are naturally nested in groups by

individual characteristics, by class, and by school. This creates the data structure of the Hierarchical Linear Model (HLM). As such, models at higher levels of the HLM structure (in this case, Level 2 and Level 3) can be formulated using various lower level units of analysis. Where nested data, such as those used in this study, occur, relationships between predictor variables and student outcomes (proficiency in oral reading fluency) can be investigated across levels. That is, student achievement can be predicted by school-level variables (principal effects) in addition to student-level variables where a hierarchical data structure exists.

This study utilized a sequential modeling technique using hierarchical data nested within students and within schools. The first phase of the analysis incorporated an analysis of variance (ANOVA) model to determine if there was sufficient variance among the student population to warrant additional modeling. If sufficient variance was found, a Level 1 model would be used to provide initial status and a resultant growth trajectory for the population of students. As it is understood that various student-level factors may affect the academic proficiency of a student, a Level 2 model would then be incorporated (providing a favorable outcome in the Level 1 model) to partition student variance to selected student-level characteristics. Whether or not a student is classified as an ethnic minority, is eligible for free or reduced lunch, or has limited English proficiency are examples of these student-level variables. If statistically-significant variation in student achievement was found in the Level 2 model, a Level 3 model would be incorporated to partition unexplained variance to school-level variables, in this case principal effects.

Analysis of Variance (ANOVA) Model

All analyses were conducted using HLM 6 (HLM; Raudenbush, Bryk & Congdon, 2001). Following appropriate HLM methodology, an ANOVA model was created to serve as a baseline for the estimation of growth trajectories in reading. More specifically, this model tested the amount of variation in ORF scores at the student level, nested in schools. ORF scores from 34,275 students from 367 *Reading First* schools were matched with the corresponding 367 principals from the total principal respondents (n=388). The purpose of this analysis was to determine if there was significant student reading achievement variation among schools and among individual students to justify

further modeling. That is, without a statistically significant amount of student reading variability at this phase of the analysis there would be no basis for additional analysis. The resulting p -value was statistically significant at $p < .001$, indicating that there was a significant amount of residual unexplained variance that can be modeled at Level 1, using time as the variable of interest (Table 11).

Table 11.
Final Estimation of Variance Components for ANOVA Model

Random Effect	<i>SD</i>	<i>p</i> -value
INTERCEPT	22.69	0.000*
Level-1	15.78	

*statistically significant $p < .001$

Level 1 Model (individual student scores over time)

A Level 1 model was created to estimate student reading achievement growth using the four instances of ORF testing (time) as the focus variable. The Level 1 model did not account for individual student characteristics (Level 2) or principal effects (Level 3). This model tested the growth of ORF scores at the student level to determine the mean initial status (oral reading fluency score) of the student population at ORF1 and the resultant mean slope of growth for the population across the remaining three instances of ORF testing (ORF2, ORF3 and ORF4). These aggregate population mean slopes were a reflection of the sum of individual student growth slopes (trajectories). In other words, Level 1 provided an understanding of where the student population started, in regard to oral reading fluency, and also how quickly the study population added correctly pronounced words per minute (WPM), what the ORF assessment measures. Table 12 shows that the initial intercept for the student population at ORF1 was statistically significant at $p < .001$ ($M=19.28$). In addition, the mean slope of growth over time for the student population was 9.37 WPM for each of the three periods of ORF assessment (from ORF1 to ORF2, ORF2 to ORF3 and ORF3 to ORF4, respectively), also statistically

significant at $p < .001$.

Table 12.
Final Estimation of Fixed Effects for Level 1

Fixed Effect	Coefficient	<i>SD</i>	<i>p</i> -value
INTERCEPT	19.28	0.33	0.000*
Time Slope	9.37	0.02	0.000*

*statistically significant $p < .001$

A resulting $p < .001$ when calculating residual variance statistics showed that there was true variance in individual student reading growth within the Level 1 model. This also indicated that there was enough remaining unexplained variance to justify further modeling of student reading growth over time, adding the effects of individual student characteristics on ORF performance (Table 13).

Table 13.
Final Estimation of Level 1 and Level 2 Variance Components

Random Effect	<i>SD</i>	<i>p</i> -value
INTERCEPT	23.48	0.000*
Level 1	10.14	

*statistically significant $p < .001$

Level 2 Model (student characteristic effects on ORF growth)

Testing the Level 1 model allowed the researcher to determine with confidence that there was significant variation in student achievement that could be potentially explained by partitioning variance to individual student characteristics. To facilitate this, a Level 2 model was created to estimate the amount of reading growth change that could

be attributed to student-level variables (i.e., SES, Gender, Ethnicity, LEP and ESE). A reference individual was developed with which to compare the reading achievement growth of the students selected for this study. For this purpose, the reference individual was a white male that was not eligible for free or reduced lunch, and that was not classified as disabled or LEP. This reference individual was used as a baseline in comparing the individual student reading growth trajectories of each student in the sample. As such, the student subgroups were comprised of students that do not display the same individual characteristics as the reference individual. For example, all students comprising the SES subgroup qualified for free/reduced lunch because the reference individual did not. Likewise, the gender subgroup contained all female students because the reference individual is male.

Table 14 displays the initial intercepts—mean scores on ORF1—by subgroup. The mean initial score on ORF for the 2004-2005 school year was 22.47 correctly pronounced words per minute (WPM), which was statistically significant at $p < .001$. This score was used to compare all other subgroup scores, so each subgroup coefficient (ORF score) was added or subtracted from the reference coefficient to get each subgroup's initial ORF score.

The SES subgroup was comprised of students eligible for free or reduced lunch. The coefficient (ORF1 score) for this group was -2.25, meaning that they pronounced 2.25 fewer words per minute than the reference individual. The Gender subgroup was comprised of females which scored 4.35 higher than the reference individual. The LEP subgroup (comprised of students classified as have limited English proficiency) scored 1.69 points lower than the reference individual, while the ESE subgroup (comprised of students classified as disabled and receiving instruction via exceptional student education) scored 0.44 points lower. All of the previously mentioned subgroup coefficients were statistically significant at $p < .001$. The Ethnicity subgroup coefficient (0.04) was not statistically significant (Table 14), thus indicating that the effects should be fixed and not used for future modeling.

Table 14.
Final Estimation of Fixed Effects for Level 2 (Initial Intercept)

Fixed Effect	Coefficient	SE	p-value
INTERCEPT (Reference)	22.47	0.43	0.000*
INTERCEPT (SES)	-2.25	0.15	0.000*
INTERCEPT (Gender)	4.35	0.26	0.000*
INTERCEPT (Ethnicity)	0.04	0.11	0.684
INTERCEPT (LEP)	-1.69	0.11	0.000*
INTERCEPT (ESE)	-0.44	0.052	0.000*

*statistically significant $p < .001$

Table 15 displays the slope—mean growth trajectories, or rate of growth in ORF—by subgroup. As can be seen, the slope of reading growth across the three subsequent instances of ORF testing (ORF2, ORF3 and ORF4) for the reference individual was 10.46 WPM. Positive coefficients reflected more WPM of growth and a slope that is steeper than the reference individual. Conversely, a negative coefficient represented fewer WPM of growth and thus a less steep slope. The SES subgroup displayed 0.69 fewer WPM of growth and the Gender subgroup displayed 1.05 more WPM of growth than the reference individual. Further, the LEP subgroup showed 0.28 fewer WPM of increase and the ESE subgroup showed 0.30 fewer WPM of growth than the reference individual. Once again, each of the previously mentioned subgroup coefficients was statistically significant at $p < .001$. The Ethnicity subgroup coefficient (0.03) was not statistically significant, identifying a lack of sufficient variability to justify further modeling at Level 3, and therefore was excluded from any further analyses.

Table 15.
Final Estimation of Fixed Effects for Level 2 (Slope)

Fixed Effect	Coefficient	SE	p-value
INTERCEPT (Reference)	10.46	0.059	0.000*
INTERCEPT (SES)	-0.69	0.027	0.000*
INTERCEPT (Gender)	1.05	0.049	0.000*
INTERCEPT (Ethnicity)	0.026	0.020	0.188
INTERCEPT (LEP)	-0.28	0.018	0.000*
INTERCEPT (ESE)	-0.30	0.01	0.000*

*statistically significant $p < .001$

A resulting $p < .001$ when calculating residual variance statistics indicated that there was significant variation in individual student reading growth, based on the selected individual characteristics. This also indicated that there was sufficient remaining unexplained variance to justify further modeling of student reading growth over time to determine the amount of reading achievement variability that could be explained at the school level. Further, an intraclass correlation was performed to determine the actual proportion of variance in student reading growth that was explained by the Level 2 model (Equation 1). It is important to note that the values used in this manual calculation are taken from HLM output reports which are not provided in this Results section. The intraclass correlation is provided in Table 16 for efficiency in reporting and does not reflect a calculation based on the other values reported in the table.

The calculation revealed that the Level 2 model accounted for approximately 55% of the total variance in reading growth among the student population, leaving approximately 45% of the total variance unaccounted for (Table 16). This proportion of variability in student ORF scores could potentially be explained by school-level variables. In this case, principal effects would be modeled to determine the amount of variance that could be explained by these variables.

Equation 1.

Intraclass Correlation Formula (Proportion of Variance at Level 2)

$$\tau_{\chi 00} / (\sigma^2 + \tau_{\chi 00} + \tau_{\beta 00})$$

(variance component of Level 1 / variance component of Level 1 + variance component of Level 2)

Table 16.

Final Estimation of Fixed Effects for Level 1 and Level 2 Variance Components

Random Effect	SD	Proportion of Variance	
		Explained by Model	p-value
INTERCEPT 1	22.79		0.000*
Level 1	10.01		
Intraclass Correlation		.55	

*statistically significant $p < .001$

Level 3 Model (principal effects on ORF growth)

Statistical results from the previous ANOVA, Level 1 and Level 2 models justified the further examination of variables associated with student reading achievement growth, based on the statistically significant residual variance. Variables selected for the subsequent Level 3 model reflected what was considered to be appropriate principal-level activities supporting the implementation of effective reading programs.

The relationship between principal implementation of selected components of effective reading programs was modeled to determine potential direct effects on reading achievement, as well as account for additional unexplained model variance. Subgroup intercept coefficients (SES, Gender, LEP and ESE) are displayed in relation to the

reference individual's intercept estimate (22.46). Results showed that the SES subgroup (all students eligible for free/reduced lunch) scored 2.24 WPM below the reference individual at ORF1 (20.22 WPM). The Gender subgroup (female students) scored 4.35 wpm higher than the reference individual (26.81 WPM). The LEP subgroup (all students classified as LEP) scored 1.69 WPM lower than the reference individual (20.77 WPM). Finally, the ESE subgroup (all students classified as ESE) scored .44 WPM lower than the reference individual (22.02 WPM). Once again, principal effects on the Ethnicity subgroup were not estimated due to a lack of significant effects identified in the Level 2 model.

Figure 2 provides a representation of all of the student subgroups' ORF growth in relation to the reference individual. The Ethnicity subgroup is not represented due to its elimination in the Level 2 model.

Population subgroups were also displayed by correlations to PIQ domains (leadership, professional development, assessment and intervention). These results indicated no significant effect of PIQ domains on the intercept (ORF1). This could be explained by the fact that this assessment occurred prior to major principal implementation activities, as measured by the annual, year end PIQ.

Principal effects were, again, modeled on the initial student starting points (i.e., intercepts) to determine what influence, if any, principal behaviors had on student ORF1 scores (Table 17). Subgroup slope coefficients (SES, Gender, LEP and ESE) are displayed in relation to the reference individual's slope coefficient (10.46, reflecting an increase of 10.46 WPM in each assessment period), with each subgroup coefficient representing measurable changes (positive or negative) in the growth of individual student reading achievement (Table 18).

Research Questions and Hypotheses

Four sub-hypotheses for each research question were considered to reflect the potential relationships between PIQ domains and ORF growth among student subgroups.

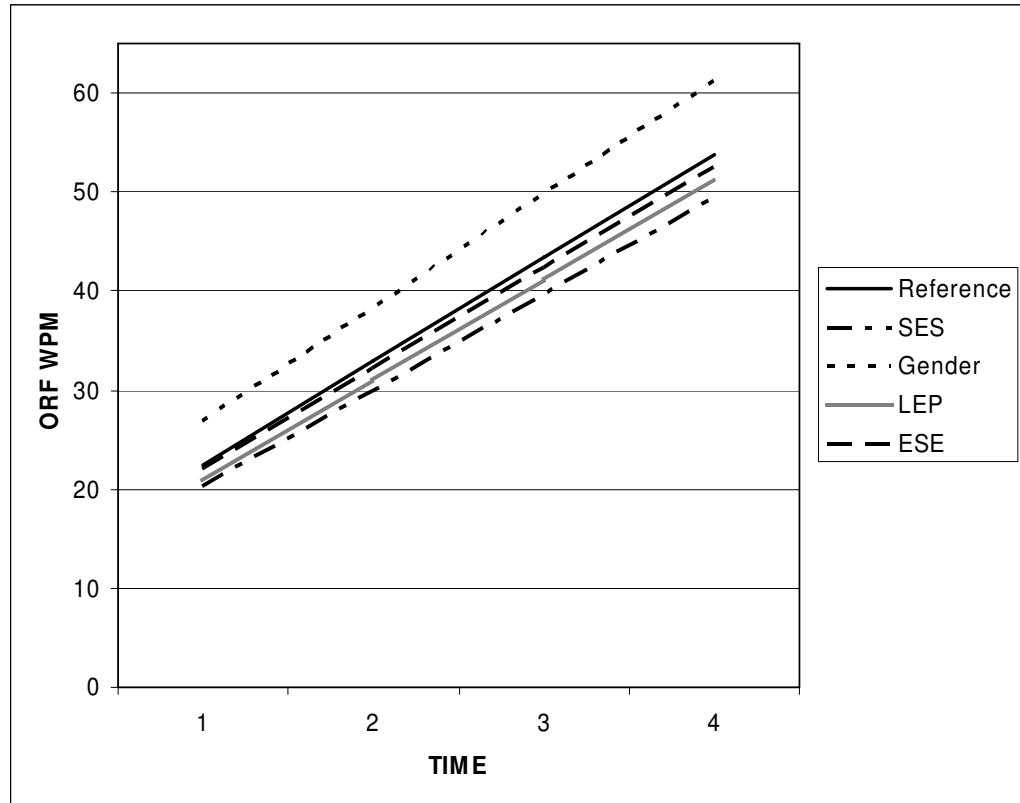


Figure 2.
Slope of ORF Growth for All Student Subgroups

Research Question #1

1. Is there a significant relationship between defined elements of principals' implementation of effective reading programs and growth in student reading achievement in the population of elementary students identified during the 2004-2005 academic year?

1.1 There is no statistically significant effect of principal behaviors in the implementation of effective reading programs on the reading achievement of students in relation to the PIQ Core Reading Program domain.

The Core Reading Program domain did not meet basic statistical requirements in the validation of the Principal Implementation Questionnaire (PIQ) and was eliminated from further analyses.

Table 17.
Final Estimation of Fixed Effects for Level 3 (Intercept)

Fixed Effect	Coefficient	SE	p-value
INTERCEPT (Reference)	22.46	0.43	0.000*
LEAD	-0.06	1.55	0.969
PROFD	0.28	1.03	0.787
ASSESS	0.73	1.37	0.593
INTERV	0.36	0.70	0.603
INTERCEPT (SES)	-2.24	0.15	0.000*
LEAD	0.45	0.57	0.454
PROFD	-0.30	0.37	0.492
ASSESS	0.33	0.50	0.545
INTERV	0.32	0.25	0.197
INTERCEPT (Gender)	4.35	0.27	0.000*
LEAD	0.45	0.92	0.623
PROFD	-0.30	0.64	0.632
ASSESS	0.33	0.83	0.691
INTERV	0.14	0.43	0.737
INTERCEPT (LEP)	-1.69	0.11	0.000*
LEAD	0.02	0.37	0.948
PROFD	-0.23	0.25	0.371
ASSESS	0.43	0.33	0.187
INTERV	0.03	0.17	0.875
INTERCEPT (ESE)	-0.44	0.05	0.000*
LEAD	-0.10	0.19	0.627
PROFD	0.04	0.13	0.731
ASSESS	-0.11	0.17	0.494
INTERV	-0.16	0.083	0.061

*statistically significant $p < .001$

1.2 There is no statistically significant effect of principal behaviors in the implementation of effective reading programs on the reading achievement of students in relation to the PIQ Leadership domain.

There was no statistically significant effect of leadership (LEAD) on reading achievement growth across the four instances of ORF assessment ($p = 0.935$). Thus, with regard to leadership, the researcher failed to reject null hypothesis 1.2.

1.3 There is no statistically significant effect of principal behaviors in the

implementation of effective reading programs on the reading achievement of students in relation to the PIQ Professional Development domain.

There was no statistically significant effect of professional development (PROFD) on reading achievement growth across the four instances of ORF assessment ($p = 0.989$). Thus, with regard to professional development, the researcher failed to reject null hypothesis 1.3.

1.4 There is no statistically significant effect of principal behaviors in the implementation of effective reading programs on the reading achievement of students in relation to the PIQ Assessment domain.

There was no statistically significant effect of assessment (ASSESS) on reading achievement growth across the four instances of ORF assessment ($p = 0.458$). Thus, with regard to assessment, the researcher failed to reject null hypothesis 1.4.

1.5 There is no statistically significant effect of principal behaviors in the implementation of effective reading programs on the reading achievement of students in relation to the PIQ Intervention domain.

There was a statistically significant effect of intervention (INTERV) on reading achievement growth across the four instances of ORF assessment ($p = 0.008$). Thus, with regard to intervention, the researcher rejected null hypothesis 1.5.

Research Question #2

Is there a significant difference between the reading achievement of first grade students of low socio-economic status and students of higher socio-economic status in relation to principal effects?

2.1 There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to the PIQ Core Reading Program domain.

The Core Reading Program domain did not meet basic statistical requirements in the validation of the Principal Implementation Questionnaire (PIQ) and was eliminated from further analyses.

Table 18.

Final Estimation of Fixed Effects for Level 3 (Slope)

Fixed Effect	Coefficient	SE	p-value
SLOPE (Reference)	10.46	0.06	0.000*
LEAD	0.02	0.22	0.935
PROFD	0.00	0.14	0.989
ASSESS	0.14	0.19	0.458
INTERV	0.25	0.10	0.008**
SLOPE (SES)	-0.69	0.03	0.000*
LEAD	0.27	0.10	0.009**
PROFD	-0.11	0.07	0.086
ASSESS	-0.19	0.09	0.036***
INTERV	-0.06	0.04	0.177
SLOPE (Gender)	1.04	0.05	0.000*
LEAD	-0.49	0.17	0.004**
PROFD	0.24	0.12	0.037***
ASSESS	0.40	0.15	0.010**
INTERV	0.06	0.08	0.425
SLOPE (LEP)	-0.27	0.02	0.000*
LEAD	0.20	0.07	0.003**
PROFD	-0.15	0.04	0.001*
ASSESS	-0.07	0.06	0.185
INTERV	0.01	0.03	0.741
SLOPE (ESE)	-0.30	0.01	0.000*
LEAD	-0.13	0.03	0.000*
PROFD	0.04	0.02	0.088
ASSESS	0.02	0.03	0.561
INTERV	-0.01	0.02	0.648

*statistically significant at $p < .001$, ** statistically significant at $p < .01$, *** statistically significant at $p < .05$

2.2 There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to the PIQ Leadership domain.

There was a statistically significant effect of LEAD on reading achievement growth across the four instances of ORF assessment ($p = 0.009$). Thus, with regard to leadership, the researcher rejected null hypothesis 2.2.

2.3 There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to the PIQ Professional Development domain.

There was no statistically significant effect of PROFD on reading achievement growth across the four instances of ORF assessment ($p = 0.086$). Thus, with regard to professional development, the researcher failed to reject null hypothesis 2.3.

2.4 There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to the PIQ Assessment domain.

There was a statistically significant effect of ASSESS on reading achievement growth across the four instances of ORF assessment ($p = 0.036$). Thus, with regard to assessment, the researcher rejected null hypothesis 2.3.

2.5 There is no statistically significant difference in the growth trajectories of students of low socio-economic status and students of higher socio-economic status in relation to the PIQ Intervention domain.

There was no statistically significant effect of INTERV on reading achievement growth across the four instances of ORF assessment ($p = 0.177$). Thus, with regard to intervention, the researcher failed to reject null hypothesis 2.5.

Research Question #3

Is there a significant difference between reading achievement of first grade students of different gender in relation to principal effects?

3.1 There is no statistically significant difference in the growth trajectories of male and female students in relation to the PIQ Core Reading Program domain.

The Core Reading Program domain did not meet basic statistical requirements in the validation of the Principal Implementation Questionnaire (PIQ) and was eliminated from further analyses.

3.2 There is no statistically significant difference in the growth trajectories of male and female students in relation to the PIQ Leadership domain.

There was a statistically significant effect of LEAD on reading achievement growth across the four instances of ORF assessment ($p = 0.004$). Thus, with regard to leadership, the researcher rejected null hypothesis 3.2.

3.3 There is no statistically significant difference in the growth trajectories of male and female students in relation to the PIQ Professional Development domain.

There was a statistically significant effect of PROFD on reading achievement growth across the four instances of ORF assessment ($p = 0.037$). Thus, with regard to professional development, the researcher rejected null hypothesis 3.3.

3.4 There is no statistically significant difference in the growth trajectories of male and female students in relation to the PIQ Assessment domain.

There was a statistically significant effect of ASSESS on reading achievement growth across the four instances of ORF assessment ($p = 0.010$). Thus, with regard to assessment, the researcher rejected null hypothesis 3.4.

3.5 There is no statistically significant difference in the growth trajectories of male and female students in relation to the PIQ Intervention domain.

There was no statistically significant effect of INTERV on reading achievement growth across the four instances of ORF assessment ($p = 0.425$). Thus, with regard to intervention, the researcher failed to reject null hypothesis 3.5.

Research Question #4

Is there a significant difference between reading achievement of first grade ethnic minority and non-minority students in relation to principal effects?

The Ethnicity subgroup was eliminated from analyses based on insignificant effects identified in the Level 2 model. Thus, the research question could not be considered in the Level 3 model.

Research Question #5

Is there a significant difference between the reading achievement of first grade students with limited English proficiency and students that are fluent in English in relation to principal effects?

5.1 There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to the PIQ Core Reading Program domain.

The Core Reading Program domain did not meet basic statistical requirements in the validation of the Principal Implementation Questionnaire (PIQ) and was eliminated from further analyses.

5.2 There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to the PIQ Leadership domain.

There was a statistically significant effect of LEAD on reading achievement growth across the four instances of ORF assessment ($p = 0.003$). Thus, with regard to leadership, the researcher rejected null hypothesis 5.2.

5.3 There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to the PIQ Professional Development domain.

There was a statistically significant effect of PROFD on reading achievement growth across the four instances of ORF assessment ($p = 0.001$). Thus, with regard to professional development, the researcher rejected null hypothesis 5.3.

5.4 There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to the PIQ Assessment domain.

There was no statistically significant effect of ASSESS on reading achievement growth across the four instances of ORF assessment ($p = 0.185$). Thus, with regard to assessment, the researcher failed to reject null hypothesis 5.4.

5.5 There is no statistically significant difference in the growth trajectories of students with limited English proficiency and students that are fluent in English in relation to the PIQ Intervention domain.

There was no statistically significant effect of INTERV on reading achievement growth across the four instances of ORF assessment ($p = 0.741$). Thus, with regard to intervention, the researcher failed to reject null hypothesis 5.5.

Research Question #6

Is there a significant difference between reading achievement of first grade students that qualify for exceptional student education (ESE) and students that have no disability classification in relation to principal effects?

6.1 There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to the PIQ Core Reading Program domain.

The Core Reading Program domain did not meet basic statistical requirements in the validation of the Principal Implementation Questionnaire (PIQ) and was eliminated from further analyses.

6.2 There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to the PIQ Leadership domain.

There was a statistically significant effect of LEAD on reading achievement growth across the four instances of ORF assessment ($p = 0.000$). Thus, with regard to leadership, the researcher rejected null hypothesis 6.2.

6.3 There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to the PIQ Professional Development domain.

There was no statistically significant effect of PROFD on reading achievement growth across the four instances of ORF assessment ($p = 0.088$). Thus, with regard to professional development, the researcher failed to reject null hypothesis 6.3.

6.4 There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to the PIQ Assessment domain.

There was no statistically significant effect of ASSESS on reading achievement growth across the four instances of ORF assessment ($p = 0.561$). Thus, with regard assessment, the researcher failed to reject null hypothesis 6.4.

6.5 There is no statistically significant difference in the growth trajectories of students with disabilities and students that have no disabilities in relation to the PIQ Intervention domain.

There was no statistically significant effect of INTERV on reading achievement growth across the four instances of ORF assessment ($p = 0.648$). Thus, with regard intervention, the researcher failed to reject null hypothesis 6.5.

A summary of the relationships among aggregate principal responses on PIQ domains and the selected student subgroups is provided in Table 19.

Table 19.
HLM Interactions by PIQ Domain and Student Group

PIQ Domain	Student Sub-group				
	Total Population	SES	Gender	LEP	ESE
Leadership (LEAD)		X	X*	X	X*
Professional Development (PROFD)			X	X*	
Assessment (ASSESS)		X*	X		
Intervention (INTERV)	X				

X = Statistically significant relationship
*Negative relationship

Another intraclass correlation was performed to determine the actual proportion of variance in student reading growth that was explained by the Level 3 model (Equation 2). The calculation revealed no measurable difference in the amount of student achievement variability explained by the Level 2 model (individual student characteristics) and that explained by the Level 3 model (principal effects). A total of 55% of the variance in ORF growth was explained when combining the effects of Level 2 and Level 3, the same amount explained before Level 3 (principal effects) were modeled (Table 20). That is, a significant proportion of student reading variance (approximately 45%) remained unexplained after principal effects were modeled.

Equation 2.

Intraclass Correlation Formula (Proportion of Variance at Level 3)

$$\tau_{\beta 00} / (\sigma^2 + \tau_{\gamma 00} + \tau_{\beta 00})$$

(variance component of Level 2 / variance component of Level 2 + variance component of Level 3)

Table 20.
Final Estimation of Fixed Effects for Level 2 and Level 3 Variance Components

Random Effect	<i>SD</i>	Proportion of Variance Explained by Model	<i>p</i> -value
INTERCEPT 1	22.78		0.000*
Level 1	10.01		
Intraclass Correlation		.55	

*statistically significant $p < .001$

CHAPTER 5

DISCUSSION AND CONCLUSIONS

This chapter discusses and offers explanations of the findings reported in Chapter Four, provides reasonable conclusions of the overall investigation and makes recommendations for practice and future research.

3-Level Growth Curve Model

HLM growth models allow the researcher to compare multiple measurements (e.g., oral reading fluency [ORF] scores) over time for individuals whose data are nested in naturally occurring settings (e.g., students in schools). At Level 1, each student's reading achievement is represented by a growth trajectory that is dependent on student-level characteristics (Level 2) which may be affected by school-level variables (Level 3). The test scores contained in Level 1 serve as the basis for the unique individual growth trajectories modeled in later stages of the hierarchical model. The applications of a growth curve model utilized to answer the research questions and test hypotheses in this study include (a) estimating the mean growth trajectory and the variation surrounding it, (b) estimating the correlation between initial status and rate of change, (c) modeling student-level predictors of initial status and change and (d) modeling school (principal)-level predictors of student status and change.

Statistical findings from the ANOVA model (see Table 11), the Level 1 Model (see Tables 12 and 13) and the Level 2 Model (see Tables 14, 15 and 16) provided significant statistical rationale to justify the Level 3 Model (see Tables 17 and 18). Subsequently, this model tested the effect of selected principal-level variables on student reading achievement growth as measured by ORF.

Principal Effects on ORF Growth

The following relationships were identified using the Level 3 Model (Tables 17

and 18), and are discussed to explain the practical utility of the findings.

Principal Effects on ORF Growth in Relation to the Total Population

Across the student population, including all subgroups, the Level 3 Model revealed a statistically significant relationship between the Intervention domain (INTERV) and student proficiency in oral reading fluency over time (Table 18). Therefore, for every one-point increase in the PIQ score on the INTERV domain, the slope of oral reading fluency achievement should increase by 0.25 words per minute (WPM) during each of the three assessment periods (ORF1 to ORF2, ORF2 to ORF3 and ORF3 to ORF4) for a total expected 0.75 WPM increase in relation to the reference individual (Figure 3).

The data reflect the potential for an increased principal focus on the implementation of effective reading intervention programs to result in student gains in oral reading fluency. Although the expected increase is small, it is nonetheless a noteworthy finding insomuch as it is a measurable direct effect of principal implementation behaviors on student reading achievement.

Principal Effects on ORF Growth in Relation to SES

Level 3 modeling identified a statistically significant relationship between the Leadership domain (LEAD) and student growth in oral reading fluency among students with lower socio-economic status (SES). Therefore, for this sample, every one-point increase in the PIQ score on the LEAD domain should result in the slope for achievement in oral reading fluency increasing by 0.27 WPM during each of the three periods of assessment (ORF1 to ORF2, ORF2 to ORF3 and ORF3 to ORF4; see Table 18). Thus, students of lower SES were shown to respond positively to principal behaviors in leadership, as measured by the PIQ. This finding directs researchers and policymakers to look more closely at the interaction of leadership behaviors and the reading achievement of disadvantaged students. The increase in the slope of growth as related to the LEAD domain is reflected in Figure 4.

A statistically significant relationship was also found between the Assessment domain (ASSESS) and student growth in oral reading fluency among students with lower socio-economic status (SES; see Table 18). However, for every one-point increase in the PIQ score on the ASSESS domain, oral reading fluency should decrease by 0.19 WPM.

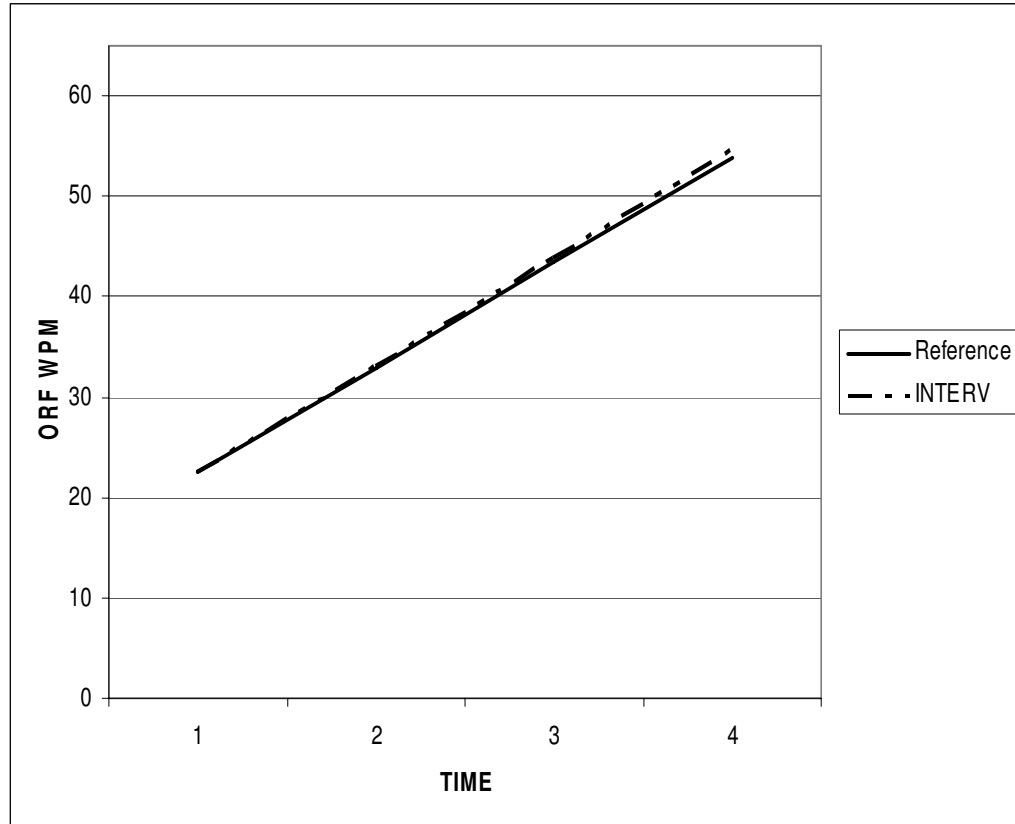


Figure 3.
Slope of ORF Growth Based On Intervention Responses for Total Population

in each of the three ORF assessment periods, for a total decrease of 0.57 WPM. This result runs counter to previous research (e.g., Waters, Marzano & McNulty, 2003), and the researcher did not expect that an increase in the implementation of reading assessment practices would yield a negative growth trajectory in relation to ORF. The confounding results suggest that the PIQ may not have captured the critical interactions of assessment on disadvantaged students and this possibility should be scrutinized more deeply. This linear relationship is illustrated in Figure 5.

Principal Effects on ORF Growth in Relation to Gender

Results identified a statistically significant relationship between the Leadership domain (LEAD) and student growth in oral reading fluency when comparing males to females (i.e., gender). Therefore, for this sample, every one-point increase in the PIQ

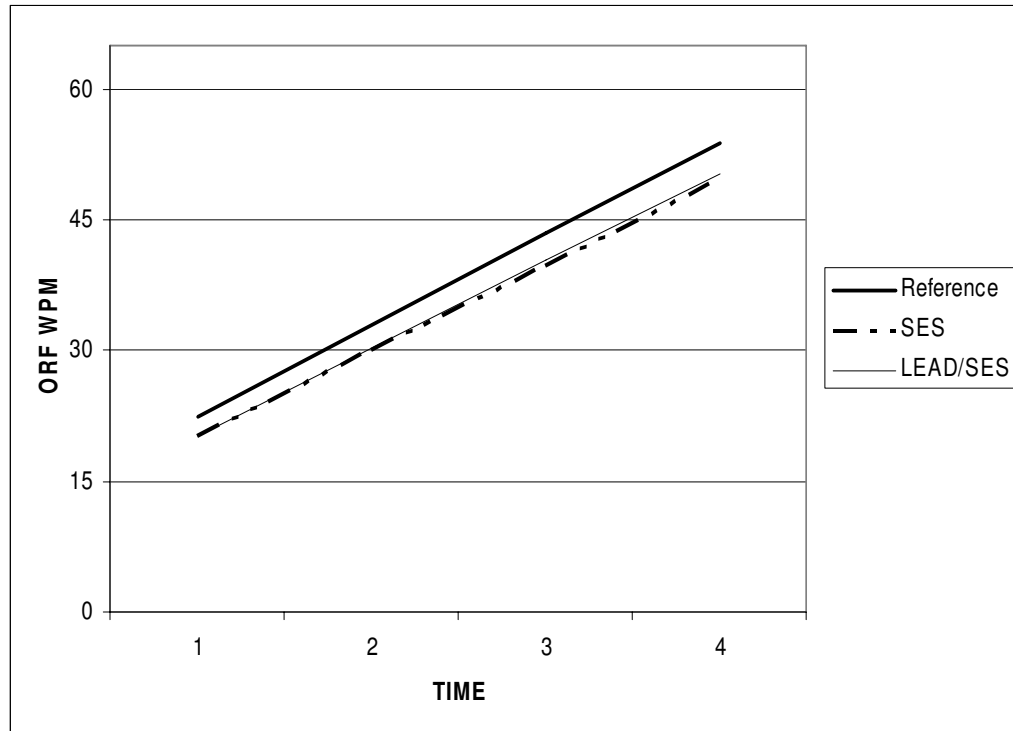


Figure 4.
Slope of ORF Growth Based On Leadership Responses for SES Subgroup

score on the LEAD domain should result in a predicted ORF decrease in female students of 0.49 WPM in each of the three assessment periods (see Figure 6). While the female students responded negatively to the implementation of leadership components measured by the PIQ, male students did increase the slope of ORF growth over the reference individual. This finding, while somewhat confounding in that it identified opposite effects for male and female students, still established a positive relationship between LEAD responses and growth in ORF and is important for that reason. A statistically significant relationship was also found between the Professional Development domain (PROFD) and student growth in oral reading fluency among male and female students (see Table 18). Results indicated that for every one-point increase in the PIQ score on the ASSESS domain, oral reading fluency is estimated to increase by 0.24 WPM per assessment period from ORF1 to ORF4 for female students (a total of .72 WPM).

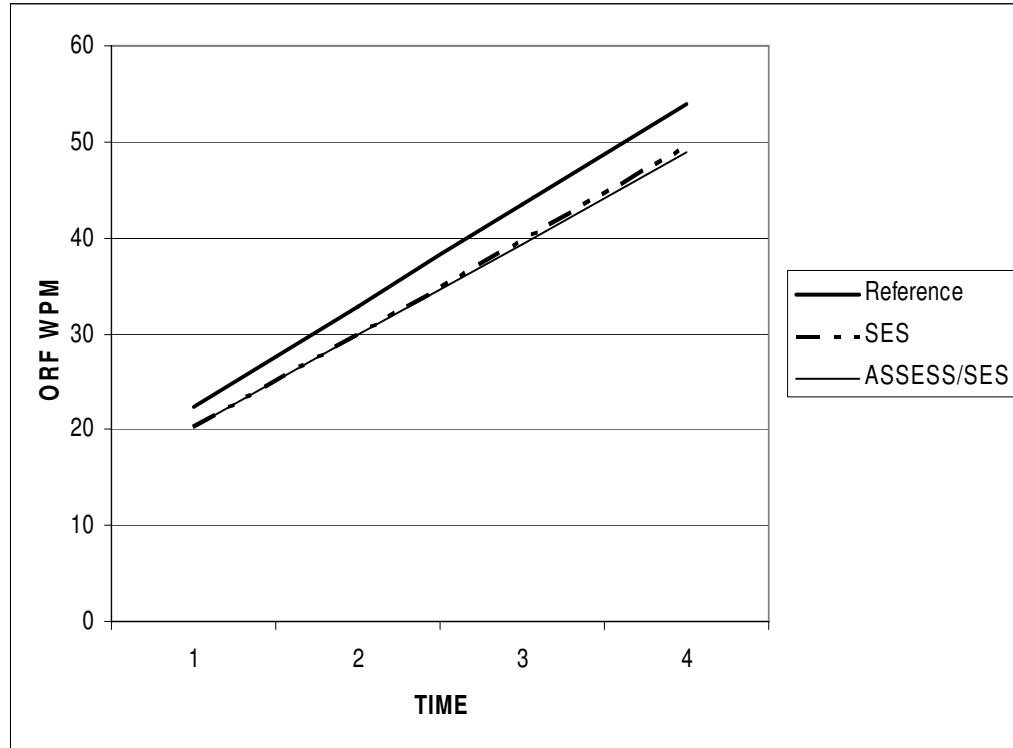


Figure 5.
Slope of ORF Growth Based On Assessment Responses for SES Subgroup

This finding is mirrored for male students as reflected in the reference trajectory (Figure 7). All students were shown to benefit from higher PIQ responses associated with professional development. Once again, while the results are sparse, they are statistically significant and noteworthy considering that these are *direct* effects of principals on student reading achievement, and not indirect effects as is commonly cited in the contemporary literature. The Assessment domain (ASSESS) also displayed a statistically-significant relationship to student growth in ORF over time (see Table 18). For every one-point increase in the PIQ score on the ASSESS domain within this sample, oral reading fluency should increase by 0.40 WPM in each of the three assessment periods, for a total of 1.20 WPM. More favorable (higher) scores on the PIQ ASSESS domain were shown to benefit all students, particularly female students (Figure 8). This relationship should be considered when developing state, district and school policies on

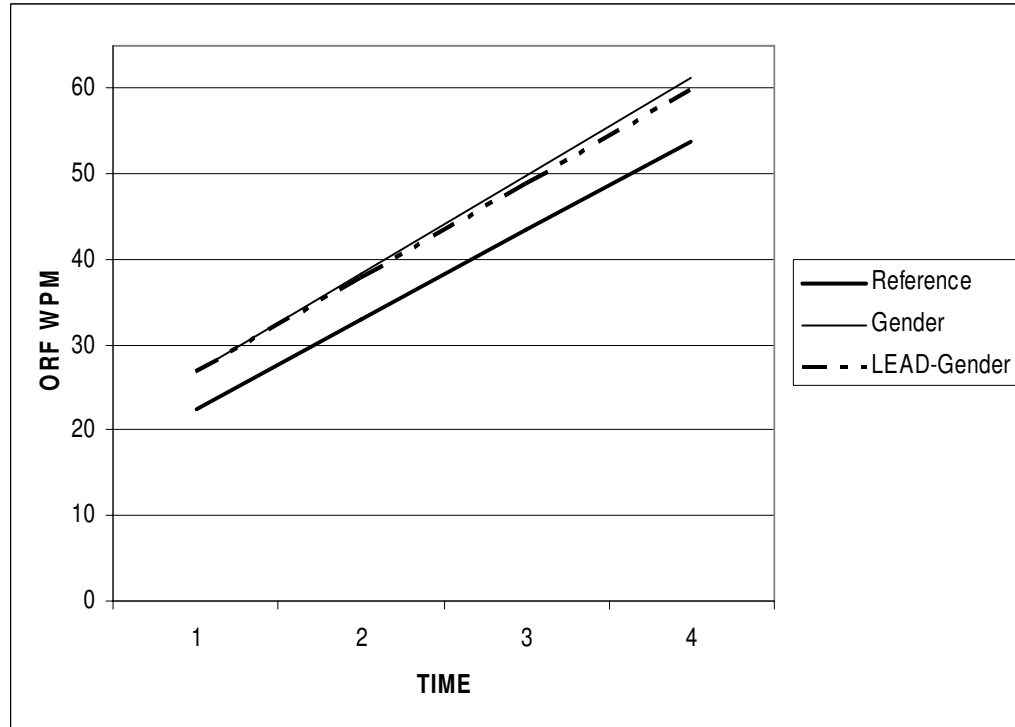


Figure 6.
Slope of ORF Growth Based On Leadership Responses for Gender Subgroup

the use of assessment protocol and the implementation of assessment plans that support reading programs for first grade students in similar populations.

Principal Effects on ORF Growth in Relation to Students with Limited English Proficiency

Level 3 modeling identified a statistically significant relationship between PIQ LEAD responses and ORF growth for students with limited English proficiency (LEP). Results indicated that for every one-point increase in the LEAD domain, ORF growth should increase by 0.19 WPM (Table 18). Leadership activities measured by the PIQ were shown to positively influence the ORF growth trajectories of students with limited English proficiency. As is displayed, the LEP subgroup growth trajectory is comparable with the reference individual, but these students begin lower and grow at a slower rate (Figure 9).

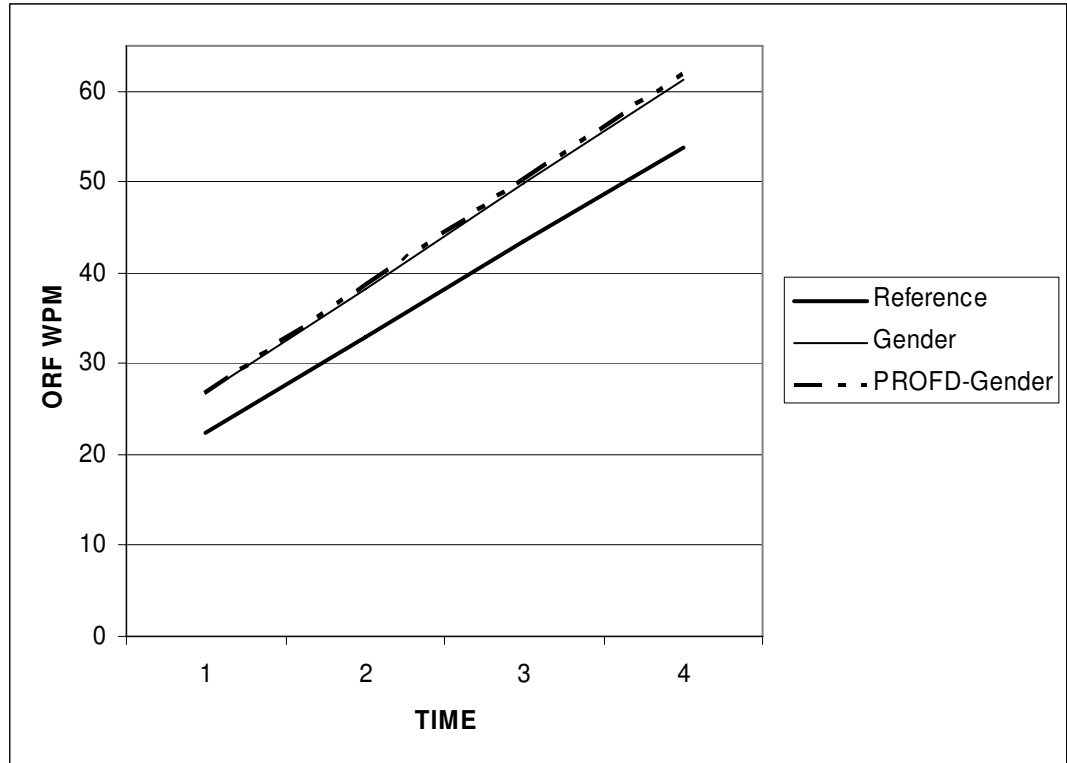


Figure 7.
Slope of ORF Growth Based On Professional Development Responses for Gender Subgroup

Statistically significant results were also found for the Professional Development domain (PROFD). However, the data showed that for every one-point increase in PROFD, it is predicted that ORF would decrease by 0.15 WPM for LEP students. The inherent limitations associated with measuring ORF among students who may not fully comprehend the language in which they are being tested may be an explanatory factor of this negative relationship (Figure 10). That is, students with limited English proficiency were asked to read English words during ORF assessments and may not have been able to decode these words; however, these students may have been able to correctly pronounce the same words had they been in their native language. Thus, ORF results for this subgroup may not have been as much of a measure of oral reading fluency as it was a measure of the extent to which students have proficiency in the pronunciation of the English language.

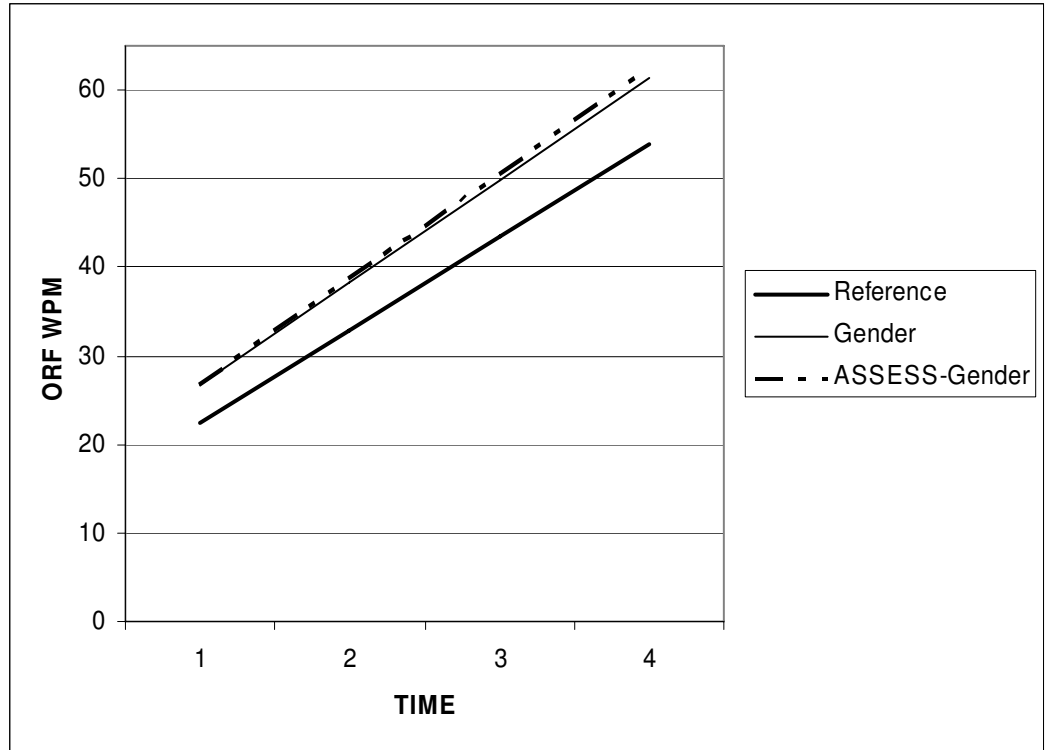


Figure 8.
Slope of ORF Growth Based On Assessment Responses for Gender Subgroup

Principal Effects on ORF Growth in Relation to Students Qualifying for Exceptional Student Education

A statistically significant relationship was found between PIQ LEAD responses and ORF growth for students with disabilities (ESE). The data showed that for every one-point increase in PROFD, it is predicted that ORF would decrease by 0.13 WPM (Figure 11). Therefore, in this population, it was found that ESE students responded negatively to principal behaviors measured by the LEAD domain with regard to achievement in oral reading fluency. While this finding is counterintuitive, the range of differences in performance and aptitude among students in 19 disability classifications and their non-disabled peers could potentially account for the negative relationship between LEAD and growth in oral reading fluency for students with disabilities.

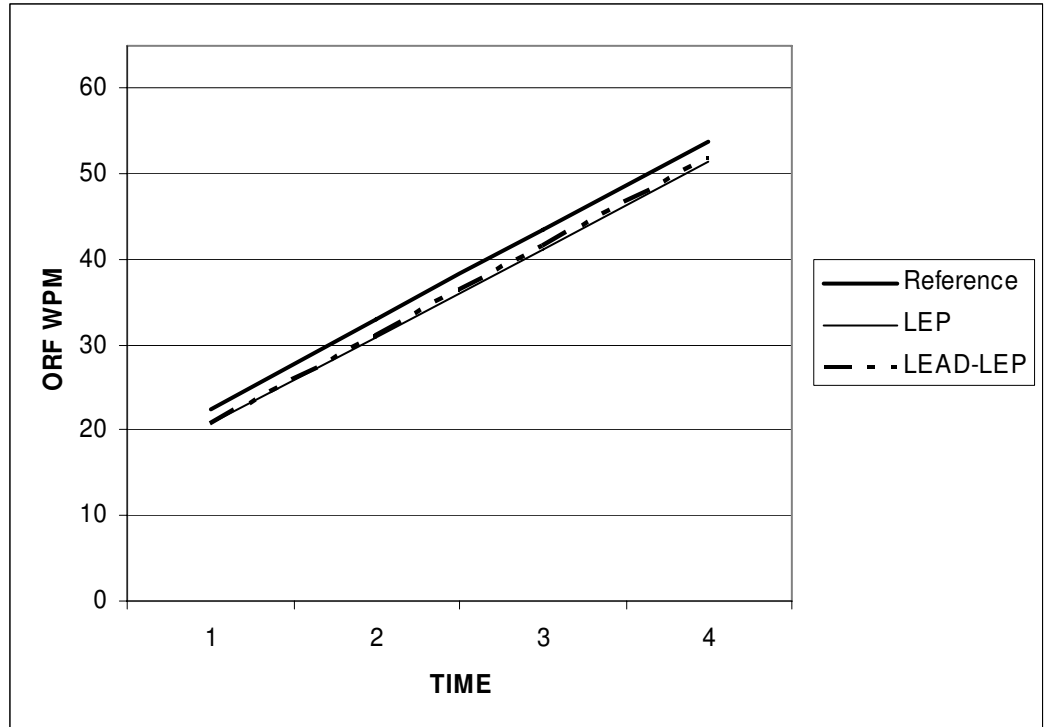


Figure 9.
Slope of ORF Growth Based On Leadership Responses for LEP Subgroup

Policymakers and researchers should account for the inherent differences among these student classifications when considering future research or practice regarding this student subgroup.

Effect Size

While the separated effects of the principal behaviors measured by the PIQ domains on the reading achievement of student subgroups were shown to be sparse, the cumulative effects of combined PIQ domain responses proved to be more substantial. For example, female students were shown to have gained 5.04 WPM from ORF1 to ORF4 when the combined effects of Professional Development and Assessment were taken into account (Table 18). This increase was .54 standard deviations higher than the

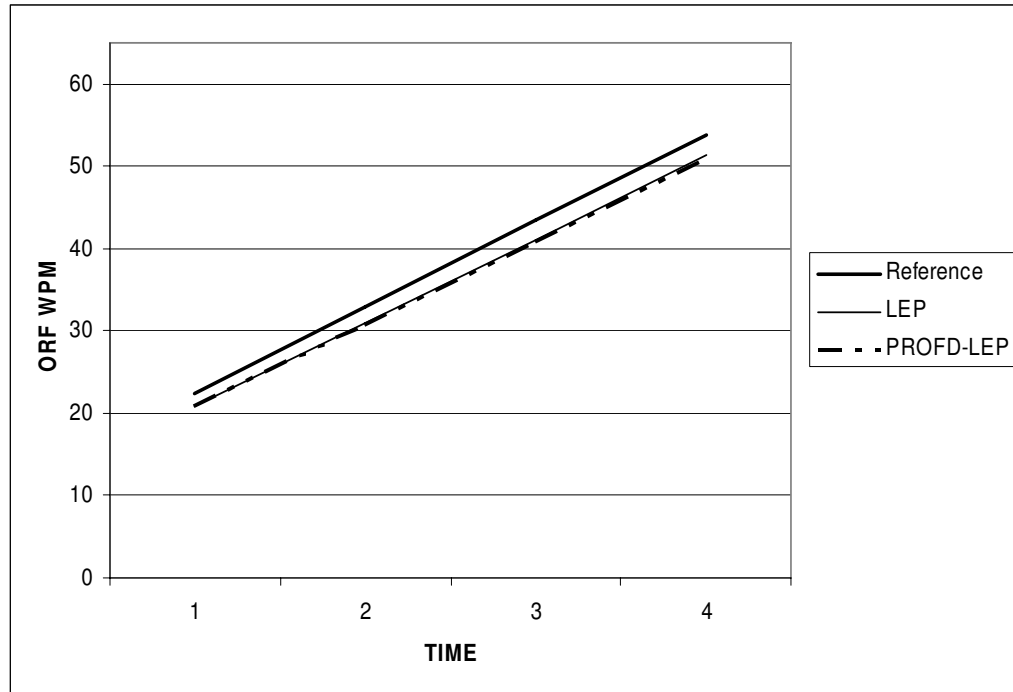


Figure 10.
Slope of ORF Growth Based On Professional Development Responses for LEP Subgroup

reference individual, due to principal effects, and resulted in an effect size of $d = .54$. This moderate effect size is the true practical importance of these findings and constitutes a noteworthy finding in that it is statistically significant and has practical utility. Conversely, the separate effects of individual PIQ domains proved to be statistically significant, but lacked robust practical significance.

Conclusions and Recommendations for Future Research

This exploratory study was narrowly designed to focus on investigating the effects of principal implementation behaviors, specifically those influencing early reading programs, on student growth in oral reading fluency. The following discussion and recommendations for future research are based on the findings of this study.

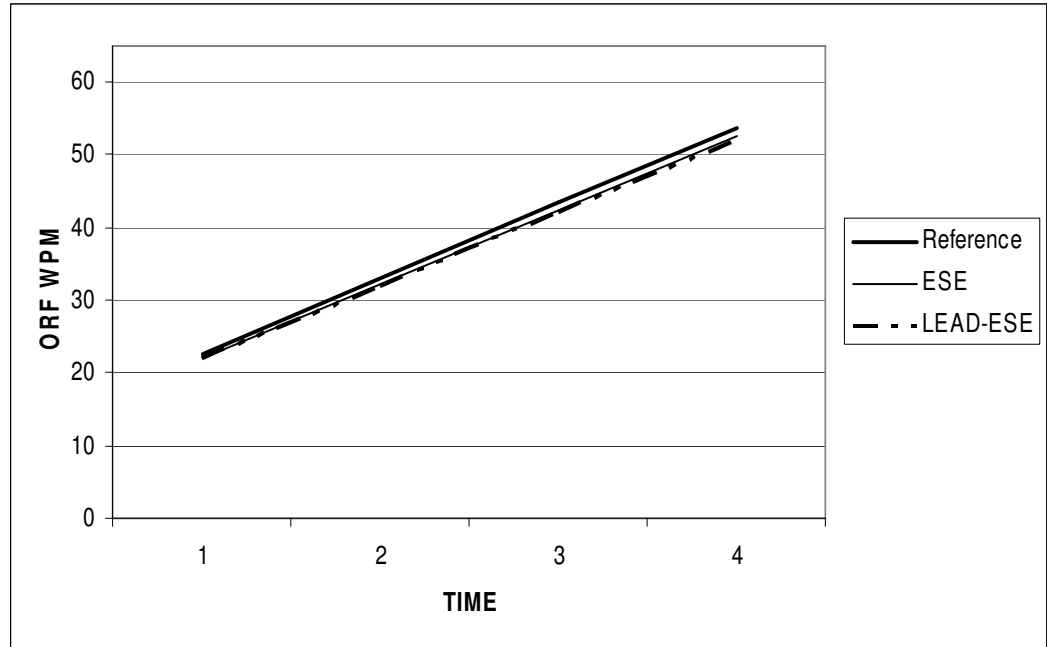


Figure 11.
Slope of ORF Growth Based On Leadership Responses for ESE Subgroup

Self-report Limitations

An overarching component of this study that impacts the understanding and discussion of the conclusions is that the Principal Implementation Questionnaire (PIQ) is a self-report instrument. This format, while commonly used in research, can be problematic inasmuch as the resulting data cannot be corroborated and is open to questions of accuracy. This fundamental challenge to the internal validity of an instrument was clearly an issue in this study, based on the design of the PIQ. In addition, the PIQ was not an anonymous survey and, thus, principal’s responses on their reading program implementation practices and perceptions could be attributed to them by the overseeing state agency (Florida Center for Reading Research). Participation in the survey was also tied to receiving *Reading First* funding, which compounded the potential that the PIQ was perceived as a high-stakes assessment. As such, the results of the PIQ have to be analyzed with the knowledge that principals may have been influenced by a “halo effect”, that is principals may have answered more positively on PIQ items in order to inflate the view of the effectiveness of reading program implementation in their

schools. Future research in this area should attempt to triangulate principal self-report data by collecting stakeholder evaluations of principal practices (e.g., teachers and/or reading coaches) and by conducting direct observations of principal implementation practices.

Direct Effects of Principal Implementation Practices

As was stated in Chapter Two, previous research has identified that principal effects are primarily indirect, as they are typically mediated by other variables more proximal to the student level. Moreover, student achievement effect sizes in relation to principal leadership have generally proven to be small. These proportions of student-level variance are, however, practically (and statistically) significant in many cases. The importance of these findings is also amplified when considered in light of the relatively small proportion of individual student achievement variance that can be attributed to endogenous variables, as is reflected in the body of research.

The findings of this study mirror the findings of other studies in regard to the proportion of direct effects and the magnitude of these effects on student achievement. Although statistically significant relationships were found between principal implementation practices as measured by the PIQ and student growth in ORF, an insignificant proportion of actual reading achievement variance could be attributed to the selected principal-level variables. In other words, a significant proportion of student reading variance (approximately 45%; Table 20) could potentially be explained by other school-level variables not captured by this model. This study reflects an absence of a notable proportion of direct effects of principal implementation behaviors and student achievement in reading. Based on the findings of previous research, this is to be expected, as principals do not have a major role in the actual classroom instruction of students. Thus, further research is needed to determine the appropriate variables that may account for this residual variance. For example, future research models should investigate the indirect effects of principal implementation practices on classroom-level variables (teacher effects) on student reading achievement. In addition, parent influences on student achievement should be investigated in this population to explain residual variance.

Implications for Instructional Leadership

The review of literature that informed this study identified many aspects of instructional leadership that can be critically important to principals in the attempt to improve student achievement. Certainly, a secondary purpose of this type of study is to determine how statistically significant results can be converted into practically important concepts and strategies that principals can use to maximize student achievement in their schools. The current results provide opportunities for researchers and practitioners to extend the statistical findings to this end.

At the conclusion of this study, it was determined that positive, statistically significant relationships existed between selected principal behaviors and increased growth in oral reading fluency. In some respects, the findings were similar to those produced by other studies. Several areas of principal behavior that have previously been identified as impacting student achievement also emerged in this study. For example, considering the 21 principal behavior constructs identified by Waters, Marzano and McNulty (2003), the emergent constructs apparent in the current study are: (1) resources; (2) knowledge of curriculum, instruction and assessment; (3) communication; (4) input; (5) ideals and beliefs; and (6) monitoring and evaluation. As such, it is recommended that these principal behaviors be further investigated to emphasize the practical value and potential implications for professional development and technical assistance. For example, these findings should be used to develop professional development opportunities tailored to teach these skills (e.g., effective implementation of reading interventions for struggling readers across all subgroups, use of leadership strategies captured by the PIQ in relation to students with limited English proficiency, etc.) to principals.

Researchers should now consider each of the positive relationships identified in this study (Tables 18 and 19) to develop a more comprehensive understanding of their unique relationship to student reading achievement. It is also suggested that negative relationships identified in this study (Table 18) be further investigated to examine these counterintuitive findings. Of particular concern are the negative relationships identified between leadership behaviors measured in the Leadership domain of the PIQ and the ORF growth of both female and ESE students (Table 18). To address these results, the

items comprising the Leadership domain should be examined to determine their appropriateness. As the PIQ was revised using an inductive process from a previous survey in which this was an existing construct, it is suggested that the behaviors reflective of this domain are potentially inappropriate or incomplete to capture positive relationships associated with this particular student population. Perhaps a revision of the Leadership domain, using more traditional behaviors associated with leadership (e.g., knowledge of curriculum, instruction and assessment, focus, communication, change agent role, or monitoring and evaluation; Waters, Marzano and McNulty, 2003), could result in positive linear relationships in regard to these student subgroups.

Future research has the potential to discriminate between students with differing classifications of disabilities in relation to principal effects on reading. For example, studies in this area could focus on students with a disability classification of speech impaired (SI), language impaired (LI) and/or specific learning disabilities (SLD) within the ESE subgroup. This type of analysis could provide additional understanding of the impact of principal implementation behaviors on discrete student subgroups.

The methodology used in this study could be replicated using other grade levels. While first grade was an appropriate and important starting point from which to begin the investigation of the effects of principal implementation behaviors and student reading achievement, the question remains: How sensitive are other early grades to principal effects in this area? Furthermore, additional DIBELS subtests should be investigated to determine the principal-level effects on growth in reading skills measured by these tests.

This study concluded that certain principal behaviors associated with implementing effective reading programs display a direct, linear relationship to student achievement. Although these findings challenge the popular notion that the true principal effects on student achievement are primarily—or solely—indirect, new evidence from the current data suggest that noteworthy direct effects exist. Further, while these relationships account for a small proportion of the total student achievement variability in reading, they are of sufficient magnitude to be of interest and additional investigation. Subsequent research in this area should be able to shed additional light on this relationship and build upon the findings of this study.

APPENDIX A
THE PRINCIPAL IMPLEMENTATION QUESTIONNAIRE
PAPER FORMAT

Principal Implementation Questionnaire- *Reading First* Schools
School Year 2004-2005

Explanation and Directions

This survey of school level instructional and supervision practices is to be completed by the building principal in consultation with the reading coach where necessary. It is required as part of the overall evaluation plan for *Reading First* in Florida.

Please answer all questions as completely and honestly as possible. The results from this survey will be used to identify areas in need of further technical assistance, support and professional development in the coming school years. *All responses will remain confidential and they will not be used in any way to evaluate individual principals or schools.* Additionally, a random selection of surveys will be followed up with school visits or telephone interviews.

There are a total of 105 questions. These questions only refer to activities in relationship to K-3 classrooms that are being supported by your *Reading First* grant for school year 2004-2005. Please circle the correct number, check the correct response, or fill in the appropriate information, as necessary. Once you complete the survey, please write the name of your county and school on the top of each page and fax the survey to the Florida Center for Reading Research at 850-644-9085. The deadline for faxing this survey is Wednesday, May 25th.

If you have any questions about how to respond to this survey, please contact Steve Nettles at snettl@fcr.org.

Your participation in this survey is appreciated.

Principal Implementation Questionnaire- *Reading First* Schools
School Year 2004-2005

Demographic Information

Please provide the following information:

1. Age: _____
2. Gender: _____ Male _____ Female
3. Race/Ethnicity:
 _____ White, Non-Hispanic:
 _____ Black, Non-Hispanic:
 _____ Hispanic
 _____ Asian or Pacific Islander:
 _____ American Indian or Alaskan Native:
 _____ Multiracial
4. Total Years of Education Experience: _____
5. Years Experience as an Administrator: _____
6. Years of Administrative Experience at Current School: _____
7. Highest Degree Attained: _____ Bachelor's
 _____ Master's
 _____ Specialist's
 _____ Doctorate
8. Year(s) of *Reading First* Implementation at Current School: _____ 1st full year
 _____ 2nd full year

Principal Implementation Questionnaire

	No classes in any grade	Some classes in some grades	All classes in some grades	All classes in all grades
1. The common core reading program (K-3) was used to guide instruction in:	1	2	3	4
2. All necessary core reading program materials became available before the start of the school year for:	1	2	3	4
3. The reading coach provided focused professional development on the implementation of the core reading program at least quarterly for:	1	2	3	4
4. The reading coach used a guided observation checklist provided by the publisher in observation of the core reading program for:	1	2	3	4
5. The reading coach provided direct feedback on the core reading program implementation for:	1	2	3	4
6. More than two days of initial training in the core reading program was provided by the publisher before materials were used in the classroom for:	1	2	3	4
7. More than two visits for ongoing training in the core reading program were conducted outside of the classroom in my school this year for:	1	2	3	4
8. Core reading program publisher trainers provided more than two classroom visits this year to model lessons and provide direct feedback for:	1	2	3	4
9. This past year, core reading program publishers provided ongoing, web-based distance learning opportunities for:	1	2	3	4

	Strongly Disagree	Disagree	Agree	Strongly Agree
10. Effective implementation of a core reading program was essential to improving reading achievement for the students in my school.	1	2	3	4
11. In general, the effective implementation of a core reading program is imperative to student improvement in reading.	1	2	3	4
12. The core reading program publisher has provided effective training on how to implement core reading program practices.	1	2	3	4
13. Professional development for teachers in implementing effective reading programs has been essential to improving student achievement in reading.	1	2	3	4
14. Professional development for principals in implementing effective reading programs has been essential to improving student achievement in reading.	1	2	3	4
15. I guided teachers in the use of reading assessment data at grade or intervention team meetings.	1	2	3	4
16. I provided opportunities for teachers to use reading assessment data to guide instructional decisions.	1	2	3	4
17. I monitored teachers' use of assessment data to guide reading instruction.	1	2	3	4
18. I was successful in finding individuals in our school to help classroom teachers in providing interventions to our most struggling readers in K-3.	1	2	3	4
19. We have successful intervention programs in place in our school for our most struggling readers in grades K-3.	1	2	3	4
20. There were consequences for teachers who did not implement effective reading programs in my school.	1	2	3	4
21. Special recognition was given to teachers who excelled in implementation of effective reading programs in my school.	1	2	3	4
22. The school schedule was constructed to provide sufficient time for reading intervention.	1	2	3	4
23. The school schedule was constructed to provide time for staff members other than classroom teachers to deliver reading interventions.	1	2	3	4

	Strongly Disagree	Disagree	Agree	Strongly Agree
24. The school schedule was constructed to provide time for teachers to collaborate on reading issues.	1	2	3	4
25. The school schedule was constructed to provide time for reading coaches to work with individual teachers as needed.	1	2	3	4
26. An effective reading coach was essential to improving student reading achievement in my school.	1	2	3	4
27. There were aspects of the <i>Reading First</i> program in Florida that interfered with our ability to provide effective instruction for all children.	1	2	3	4
28. Classroom observation by the principal was essential to improving student achievement in reading.	1	2	3	4
29. I actively promoted the full implementation of reading programs in my school.	1	2	3	4
30. Our assessment team was essential to overall school-wide reading performance this year.	1	2	3	4
31. We currently have an effective plan for implementation of <i>Reading First</i> Assessment.	1	2	3	4
32. A comprehensive screening process is currently in place to assess student progress in reading.	1	2	3	4
33. Training teachers in utilizing the core reading program to provide effective instruction is a high priority.	1	2	3	4
34. Training teachers in managing classroom behaviors so that students are more engaged in learning is a high priority.	1	2	3	4
35. Training teachers in utilizing assessment data to individualize instruction is a high priority.	1	2	3	4
36. Training teachers in organizing the classroom so that small group instruction can be provided to students who most need it is a high priority.	1	2	3	4
37. Training teachers in providing more effective interventions for students who are struggling the most in reading is a high priority.	1	2	3	4
38. Training teachers in consistently applying explicit instructional strategies when introducing new concepts is a high priority.	1	2	3	4

	Never	Sometimes	Frequently	Very Frequently
39. I used the PMRN to check overall student progress at my school..	1	2	3	4
40. I used the PMRN to examine the progress in specific classrooms within the school.	1	2	3	4
41. I used the PMRN to examine the progress of specific children.	1	2	3	4
42. I used the PMRN to compare the progress of my school against other schools who serve similar children.	1	2	3	4
43. Matters of administration were addressed in grade-level team meetings.	1	2	3	4
44. Common lesson planning and sharing was addressed in grade-level team meetings.	1	2	3	4
45. Student scheduling issues were addressed in grade-level team meetings.	1	2	3	4
46. Updates on school scheduling were addressed in grade-level team meetings.	1	2	3	4
47. Progress of individual student's intervention planning was addressed in grade-level team meetings.	1	2	3	4
48. Cooperative study and professional development was addressed in grade-level team meetings.	1	2	3	4
49. Assessment data for individual students was addressed in grade-level team meetings.	1	2	3	4
50. The lead teacher at each grade-level led the grade-level team meetings in K-3.	1	2	3	4
51. The reading coach led the grade-level team meetings in K-3.	1	2	3	4
52. The principal led the grade-level team meetings in K-3.	1	2	3	4
53. Other personnel led the grade-level team meetings in K-3.	1	2	3	4
54. Based on classroom observation, I provided feedback to teachers regarding instructional practices.	1	2	3	4
55. I monitored the use of data in guiding team level meetings.	1	2	3	4

	Never	Sometimes	Frequently	Very Frequently
56. I provided opportunities for the reading coach(es) to give feedback on the implementation of effective reading programs in my school.	1	2	3	4
57. I provided opportunities for teachers to give feedback on the implementation of effective reading programs in my school.	1	2	3	4
58. The reading coach participated on our school-level assessment team.	1	2	3	4
59. The school counselor participated on our school-level assessment team.	1	2	3	4
60. Paraprofessionals on staff participated on our school-level assessment team.	1	2	3	4
61. Classroom teachers participated on our school-level assessment team.	1	2	3	4
62. The school psychologist participated on our school-level assessment team.	1	2	3	4
63. Other administrative personnel participated on our school-level assessment team.	1	2	3	4
64. School assessment data were used to allocate appropriate resources for critical instructional areas.	1	2	3	4
65. Stakeholder feedback was monitored to assess the effectiveness of the school assessment plan	1	2	3	4
66. We used assessment data to provide feedback to teachers regarding instructional practices.	1	2	3	4
67. We used assessment data to determine professional development needs of staff.	1	2	3	4
68. We used assessment data to determine which intensive instructional practices were most effective.	1	2	3	4
69. We used assessment data to determine needed support from outside sources (i.e., district office, FCRR, Just Read, Florida!).	1	2	3	4
70. We used assessment data to determine growth patterns in struggling students' reading achievement.	1	2	3	4
71. We used assessment data to determine students' strengths and weaknesses in reading.	1	2	3	4

	Never	Sometimes	Frequently	Very Frequently
72. We used assessment data to plan for implementation of interventions for struggling readers.	1	2	3	4
73. Regular classroom teachers provided interventions to students who needed them during the regularly scheduled reading block.	1	2	3	4
74. ESE teachers provided interventions in the regular classrooms during the regularly scheduled reading block.	1	2	3	4
75. Resource teachers provided interventions in the regular classrooms during the regularly scheduled reading block.	1	2	3	4
76. Trained paraprofessional were used to provide interventions during the regularly scheduled reading block.	1	2	3	4
77. We used computer software to provide interventions during the regularly scheduled reading block.	1	2	3	4
78. We used AmeriCorps members to provide interventions during the regularly scheduled reading block.	1	2	3	4
79. We used school volunteers to provide interventions during the regularly scheduled reading block.	1	2	3	4
80. We used computer software to provide interventions <i>outside</i> of the regularly scheduled reading block.	1	2	3	4
81. We used ESE teachers to provide additional instruction in reading <i>outside</i> of the regularly scheduled reading time.	1	2	3	4
82. We used resource teachers to provide additional instruction in reading <i>outside</i> of the regularly scheduled reading time.	1	2	3	4
83. We used paraprofessionals to provide additional instruction in reading <i>outside</i> of the regularly scheduled reading time	1	2	3	4
84. We used AmeriCorps members to provide interventions <i>outside</i> of the regularly scheduled reading block.	1	2	3	4

	Never	Sometimes	Frequently	Very Frequently
85. We used school volunteers to provide interventions <i>outside</i> of the regularly scheduled reading block.	1	2	3	4
86. We used published intervention programs to help guide interventions for our struggling students.	1	2	3	4

Please rank how important the following resources were in providing technical assistance for implementation of an effective reading program at your school.

	Not Important	Important	Somewhat Important	Very Important
87. District Resources	1	2	3	4
88. State Dept. of Education Resources	1	2	3	4
89. Core Reading Program Publisher	1	2	3	4
90. Regional Professional Development Coordinator	1	2	3	4
91. Florida Center for Reading Research	1	2	3	4

92. Please indicate the percentage of your K-3 teachers who attended the 2004 Summer Reading Academies.

___ Less than 50% ___ 51-75% ___ 76-90% ___ 91-99% ___ 100%

93. In addition to the 2004 Summer Reading Academies and/or Reading Coach training, your teachers received:

___ No additional training ___ Less than a half-day of training ___ A half-day to one day of training ___ More than one day of training

94. Initial core reading program training was provided in the following format.

___ None was provided ___ Group instruction only ___ Group instruction with some individual follow-up ___ Group instruction with extensive follow-up

95. In addition to core reading program publisher training and/or the Summer Leadership Conference training, I received _____ on effective implementation of reading programs.

___ No additional training ___ Less than a half-day of training ___ A half-day to one day of training ___ More than one day of training

96. I received my earliest training in the use of my school's core reading program:

I did not receive training Before the start of the current school year. In the first half of the current school year. In the second half of the current school year.

97. Before the start of school, I received _____ in use of the core reading program.

No additional training Less than a half-day of training A half-day to one day of training More than one day of training

98. Please indicate how many times during the past year (on average) you visited each K-3 classroom in your school to observe reading instruction.

I did not visit individual classrooms to observe instruction
 I visited each classroom 1-2 times during the year
 I visited each classroom 3-6 times during the year
 I visited each classroom 7-10 times during the year
 I visited each classroom more than 10 times during the year

99. Please indicate how frequently you met with your reading coach to discuss the reading instruction or student progress for K-3 students at your school.

I was not able to meet with my reading coach this year
 I met with my reading coach 1-3 times during the year
 I met with my reading coach 4-6 times during the year
 I met with my reading coach 7-10 times during the year
 I met with my reading coach more than 10 times during the year

100. How frequently did your teachers hold grade-level team meetings?

Quarterly
 Monthly
 More than once a month, but less than once a week
 Weekly

101. How frequently did you attend grade-level team meetings?

I did not attend grade-level team meetings.
 I attended once or twice a year.
 I attended half of them.
 I attended all grade level team meetings.

102. Please indicate how much time is scheduled for reading instruction in your school every day

Less than 30 minutes
 30-90 minutes
 91-120 minutes
 More than 120 minutes

103. Please indicate the number of people who currently serve on your school-based assessment team.

- ___ 1-2
- ___ 3-4
- ___ 5-6
- ___ 7-8
- ___ 9-10
- ___ more than 10

104. How many K-3 teachers were new to your school this year?

105. How many teachers were new to your school this year due to teacher turnover? For example, this number should not reflect teachers added due to student population growth or class size reduction efforts.

If appropriate, what immediate intensive interventions were provided to students by a method NOT including: regular, ESE and resource teachers; trained paraprofessionals; computer software; AmeriCorps volunteers; or school volunteers?

If you used published intervention programs to help guide interventions for struggling students at your school, please list the names of the most widely utilized programs.

If computers were used to provide interventions for struggling students, please indicate the names of the three most commonly used pieces of software.

If you had difficulties implementing the *Reading First* Assessment plan this year, please describe the nature of your most significant difficulties.

If your school/district required reading assessment for students in grades K-3 other than those required by *Reading First* (other than one assessment of the FCAT in third grade) please indicate what assessments were used.

Kindergarten _____
First Grade _____
Second Grade _____
Third Grade _____

If your school received any technical assistance from you district to help implement your *Reading First* Grant, please describe the nature of the assistance you received.

If your teachers received any formal professional developments in reading beyond that provided by the Summer Academies and by your Reading Coach, please indicate who provided the professional development, the topic and approximately how many hours it involved.

Who _____
Topic _____
Number of Hours _____

If your school did NOT use a common core reading program to guide instruction in every classroom in grades K-3, please explain.

REFERENCES

- Anderson, L.W. (1985). Concerns for appropriate instrumentation in research on classroom teaching. *Evaluation in Education*, 8(2), 133-152.
- Bamburg, J. & Andrews, R. (1989). School goals, principals and achievement. *School Effectiveness and School Improvement*, 2(3), 175-191.
- Bamburg, J. & Andrews, R. (1991). School goals, principals and achievement. *School Effectiveness and School Improvement* 2(3), 175-191.
- Bateman, D. & Bateman, C. F. (2001). *A principal's guide to special education*. Arlington, VA: Council for Exceptional Children.
- Becker, G. (1964). *Human capital*. New York: National Bureau of Economic Research.
- Belfield, C. R. (2000). *Economic principals for education*. Cheltenham, UK: Edward Elgar.
- Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588–606.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238–246.
- Berends, M., Bodilly, S. & Kirby, S. (2002). New American schools, final report summary of RAND's report. *Phi Delta Kappan*, 84(2), 168-175.
- Biddle, B. (1997). Foolishness, dangerous nonsense and real correlates of state differences in achievement. *Phi Delta Kappan*, 79(1), 8-13.
- Bidwell, C. E. & Kasarda, J. D. (1975). School district organization and student achievement. *American Sociological Review*, 40(1), 55–70.
- Blatchford P., Goldstein H. & Mortimore P. (1998). Research on class size effects: a critique of methods and a way forward. *International Journal of Educational Research*, 29, 691-710.
- Blatchford, P. & Mortimore, P. (1994). The issue of class size for young children in schools: What can we learn from research? *Oxford Review of Education*, 20(4), 411–428.
- Blase, J. & Blase Jo. (2000). Effective instructional leadership: Teachers' perspectives on how principals promote teaching and learning in schools. *Journal of Educational*

Administration, 38(2), 130-41.

- Blase, J. & Kirby, P. C. (December 1992). The power of praise - A strategy for effective principals. *NASSP Bulletin*, 69-77.
- Bloom, B.S. (1976). *Human characteristics and student learning*. New York: McGraw-Hill.
- Bollen, K. A. (1990). Overall fit in covariance structure models: Two types of sample size effects. *Psychological Bulletin*, 107, 256–259.
- Bossert, S. T. (1988). School effects. In N. J. Boyan (Ed.), *Handbook of research on educational administration* (pp. 341-352). New York: Longman.
- Bossert, S. T., Dwyer, D. C., Rowan, B. & Lee, G. V. (1982). The instructional management role of the principal. *Educational Administration Quarterly*, 18(3), 34–64.
- Boyer, E. (1995). *The basic school: A community for learning*. Princeton, NJ: The Carnegie Foundation.
- Brophy, J.E. & Good, T.L. (1970). Teachers' communication of differential expectations for children's classroom performance. *Journal of Educational Psychology*, 61, 365-374.
- Brophy, J.E. & Good, T.L. (1986). Teacher behavior and student achievement. In M.L. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328-375). New York: MacMillan.
- Brookover, W. B. & Lezotte, L. (1982). *Creating effective schools*. Holmes Beach, FL: Learning Publication.
- Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
- Bryk, A. S. & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Butler, B.S. (1997). Web-based instruction (WBI): Using the world wide web to support classroom-based education: Conclusions from a multiple-case study. In B. H. Khan (Ed.), *Web-based instruction* (pp. 417-423). New Jersey: Educational Technology Publications.
- Cheng, Y.C., (1994). Principal's leadership as a critical factor for school performance: Evidence from multi-levels of primary schools. *School Effectiveness and School*

Improvement, 5(3), 299-317.

Clark, S. & Clark, D. (2001). The challenge of curricular and instructional improvement in an era of high stakes testing. *Middle School Journal*, 32(2), 52-59.

Cohen, D. K. & Hill, H. (1997). *Instructional Policy and Classroom Performance: The Mathematics Reform in California*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Cotton, K. (1996). Affective and social benefits of small-scale schooling. Eric Digest, EDO-RC-96-5. Cotton, K. (2000). *The schooling practices that matter most*. Alexandria, VA: Association for Supervision and Curriculum Development.

Cotton, K. (2003). *Principals and student achievement: What the research says*. Alexandria, VA: Association for Supervision and Curriculum Development.

Coleman, J. S., Campbell, E. Q., Hobson, C.J., McPartland, J., Mood, A. M., Weinfeld, F. D. & York, R. L. (1966). *Equality of educational opportunity*. US Government Printing Office: Washington, DC.

Cuban, L. (1988). *The managerial imperative and the practice of leadership in schools*. Albany, NY. State University of New York Press.

Darling-Hammond, L. (December, 1999). *Teacher quality and student achievement: A review of state policy evidence*. Center for the Study of Teaching and Policy, Seattle, WA: University of Washington.

Darling-Hammond, L. (2000). How teacher education matters. *Journal of Teacher Education*, 51(3), 166-173.

Darling-Hammond, L. & Youngs, P. (2002). Defining "highly qualified teachers": What does "scientifically-based research" actually tell us? *Educational Researcher*, 31(9), 13-25.

Deal, T. E. & Peterson, K. D. (1990). *The principal's role in shaping school cultures*. Washington, DC: U.S. Department of Education.

Deal, T. E. & Peterson, K.D. (1999). *Shaping school culture: The heart of leadership*. San Francisco: Jossey-Bass.

Donahue, P., Daane, M. & Grigg, W. (2003). *The nation's report card: Reading highlights of 2003*, (NCES-2004452), Washington, DC: U.S. Department of Education.

DuFour, R. & Berkey, T. (1995). The principal as staff developer. *Journal of Staff Development* 16(4).

- Durkin, D. (1993). *Teaching them to read*. (6th ed.) Boston, MA: Allyn and Bacon.
- Eccles, J.S., Midgley, C., Wigfield, A., Buchanan, C.M., Reuman, D., Flanagan, C. & MacIver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, 48(2), 90-101.
- Ediger, M. (1999). The principal and the reading curriculum. Monograph, ERIC# ED432006. Ediger, M. (2000). The school principal as leader in reading instruction. *Reading Improvement*, 37(1), 20-29.
- Edmonds, R. (1979, October). Effective schools for the urban poor. *Educational Leadership*, 37, 15-24.
- Edmonds, R. (February, 1982). *Programs of school improvement*. Paper presented at the National Invitational Conference, Research on Teaching: Implications for Practice, Warrenton, VA.
- Ellinger, K., Wright, D.E. & Hirlinger, M.W. (1995). Brains for the bucks? School revenue and student achievement in Oklahoma. *The Social Science Journal*, 32(3), 299-308.
- Elmore, R. (2000) *Building a new structure for school leadership*. Washington, DC: Albert Shanker Institute. [Online] Available at <http://www.ashankerinst.org/Downloads/building.pdf>.
- Evans, L. & Teddlie, C. (1995) Facilitating change in schools: Is there one best style? *School Effectiveness and School Improvement*, 6, 1-22.
- Fink, E. & Resnick, L. (2001, April). Developing principals as instructional leaders. *Phi Delta Kappan*, 82, 598-606.
- Finn, J. & Achilles, C. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal*, 27(3), 557-577.
- Flanagan, J.C., Tiedeman, D.V., Clemans, W.V. & Wise, L.L. (1964). *Project talent: The identification, development and utilization of human talents* (Final Report for Cooperative Research Project No. 635, U.S. Office of Education). Pittsburgh, PA: University of Pittsburgh.
- Florida Department of Education. (2002). *Reading First in Florida*. [Online] Available at http://www.justreadflorida.com/docs/reading_grant.pdf
- Fortune, J.C. & O'Neill, J.S. (1994). Production function analyses and the study of educational funding equity: A methodological critique. *Journal of Educational*

Finance, 20, 107-131.

- Freiberg, H. J. (1998). Measuring school climate: Let me count the ways. *Educational Leadership*, 56(1), 22-26.
- Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5, 239-256.
- Fullan, M. (1997). *What's worth fighting for in the principalship*. Toronto: Ontario Public School Teachers' Association.
- Fullan, M. (2003). *The moral imperative of school leadership*. Thousand Oaks, CA: Corwin.
- Fullan, M. & Hargreaves, A. (1991). *What's worth fighting for: Working together for your school*. Toronto: Ontario Public School Teachers' Federation.
- Ferguson, R. F. & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama schools. In H. Ladd (Ed.), *Holding Schools Accountable*. (pp. 265-298). Washington, DC: Brookings Institution.
- Fuchs, L.S., Fuchs, D., Hosp, M.K. & Jenkins, J.R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical and historical analysis. *Scientific Studies of Reading*, 5, 239-256.
- Gates, S., Ross, K. & Brewer, D. (2001). *Leading to reform: Educational leadership for the 21st century*. Oak Brook, IL: North Central Regional Educational Laboratory.
- Glass, G., Cahen, L., Smith, M. & Filby, N. (1989). *School class size*. Beverly Hills, CA: Sage.
- Good, R., Simmons, D. & Kame'enui, E. (2001). The importance and decision-making utility and a continuum of fluency-based indicators of foundational reading skills for third-grade high stakes outcomes. *Scientific Studies in Reading*, 5, 257-288.
- Good, R.H. & Kaminski, R.A. (Eds.). (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.
- Good, T.L. & Brophy, J.E. (1984). *Looking in classrooms* (3rd ed.). New York: Harper & Row.
- Graham, M. W. (1997, March). *School principals: Their roles and preparation*. Paper presented at the National Conference on Creating the Quality School, Oklahoma City, OK. (ERIC Document Reproduction Service No. ED411576)

- Greenwald, R., Hedges, L.V. & Laine, R.D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, 66(3), 361-396.
- Grissmer, D.W., Flanagan, A.E., Kawata, J.H. & Williamson, S. (2000) *Improving student achievement: What state NAEP test scores tell us*. [Online] Available at <http://www.rand.org/publications/MR/MR924/>.
- Gullatt, D.E. & Lofton, B.D. (1996). *Programs of assistance for the at-risk student*. (ERIC Document Reproduction Service No. ED 403 228)
- Hallinger, P. (November 2003). Leading educational change: Reflections on the practice of instructional and transformational leadership. *Cambridge Journal of Education*, 33(3), 329 – 352.
- Hallinger, P. (2000). A review of two decades of research on the principalship using the Principal Instructional Management Rating Scale. Paper presented at the annual meeting of the American Educational Research Association, Seattle, Washington.
- Hallinger, P., Bickman, L. & Davis, K. (1996). School context, principal leadership and student reading achievement. *The Elementary School Journal*, 96, 527-549.
- Hallinger, P. & Heck, R. (2002). What do you call people with visions? The role of vision, mission and goals in school leadership and improvement. In K. Leithwood & P. Hallinger (Eds.), *Second international handbook of educational leadership and administration* (pp. 9–40). Netherlands: Kluwer.
- Hallinger, P. & Heck, R. (1999). Can leadership enhance school effectiveness? In Bush, T., Bell, L., Bolam, R., Glatter, R. and Ribbins, P. (Eds.) *Educational management: Redefining theory, policy and practice*. London: Paul Chapman.
- Hallinger, P. & Heck, R. H. (1998). Exploring the principal's contribution to school effectiveness: 1980-1995. *School Effectiveness and School Improvement*, 9, 157-191.
- Hoy, W. K. & Ferguson, J. (1985, Spring). A theoretical framework and exploration of organizational effectiveness of schools. *Educational Administration Quarterly*, 21(2), 117-134.
- Hallinger, P. & Heck, R. (1996). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980-95. *Educational Administration Quarterly*, 32 (1) 5-44.
- Hallinger, P. & Leithwood, K. (1994). Introduction: Exploring the impact of principal leadership. *School Effectiveness and School Improvement*, 5, 206-218.

- Hanushek, E. A. (1981). Throwing money at schools. *Journal of Policy Analysis and Management*, 1, 19-41.
- Hanushek, E.A. (1986). The economics of schooling. *Journal of Economic Literature*, 24, 1141-1177.
- Hanushek, E.A. (1989). The impact of differential expenditures on school performance. *Educational Researcher*, 18(4), 45-51.
- Hanushek, E. A. (1996). School resources and student performance. In G. Burtless (Ed.). *Does money matter? The effect of school resources on student achievement and adult success*, 43-73. Washington, D.C.: Brookings Institution.
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational evaluation and policy analysis* 19(2), 141-64.
- Hanushek, E. A. (1999), The evidence on class size. In S. E. Mayer & P. Peterson (Eds.). *Earning and learning: How schools matter*, 131-168. Washington, DC: Brookings Institutions.
- Hartman, W.T. (1994). District spending disparities revised. *Journal of Education Finance*, 20, 88-106.
- Heck, R. H. (1992). Principals' instructional leadership and school performance: Implications for policy development. *Educational Evaluation and Policy Analysis*, 14(1), 21-34.
- Heck, R. H. (1993). School context, principal leadership and achievement: The case of secondary schools in Singapore. *The Urban Review*, 25(2), 151-166
- Heck, R.H. & Marcoulides, G.A. (1996). School culture and performance: Testing the invariance of an organizational model. *School Effectiveness and School Improvement*, 7(1), 76-95.
- Hedges, L.V., Laine, R. D. & Greenwald, R. (1994). Does money matter? A meta-analysis of the effects of differential school inputs on student outcomes. *Educational Researcher*, 23(3), 5-14.
- Herriot, R.F. & Firestone, W.A. (1984). Two images of schools as organizations: A refinement and an elaboration. *Educational Administration Quarterly*, 20(4), 41-58.
- Hoy, W. K. & Miskel, C.G. (2001). *Educational administration: Theory, research and practice (6th ed)*. New York: Mc Graw-Hill.

- Hoy, W. K., Smith, P. A. & Sweetland, S. R. (2002). The development of the organizational climate index for high schools: Its measure and relationship to faculty trust. *The High School Journal*, 86(2), 38-49.
- Hoy, W. K., Sweetland, S.R. & Smith, P.A. (2002). Toward and organizational model of achievement in high schools: The significance of collective efficacy. *Educational Administration Quarterly*, 38(1), 77-93.
- Hu, L. & Bentler, P. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling. Concepts, issues and applications* (pp. 76–99). London: Sage.
- Ingersoll, R. M. (1996). Teachers decision-making power and school conflict. *Sociology of Education*, 69(2), 159–176.
- Jencks, C., Smith, M., Ackland, H., Bane, M., Cohen, D. Gintis, H., Heyns, B. and Michelson, S. (1972). *Inequality: A reassessment of the effects of family and schooling in America*. New York: Basic Books.
- Johnson, J. F. & Asera, R. (Eds.). (1999). *Hope for urban education: A study of nine high-performing, high-poverty urban elementary schools*. Austin, TX: University of Texas, Charles A. Dana Center. [Online] Available at www.ed.gov/pubs/urbanhope
- Johnson, W. L. & Johnson, M. (1993). Validity of the quality of school life scale: A primary and second-order factor analysis. *Educational & Psychological Measurement*, 53(1), 145-153.
- Johnson, W. L. & Johnson, A. M. (1997). Assessing the validity of scores on the Charles F. Kettering Scale for the junior high school. *Educational & Psychological Measurement*, 57(5), 858-869.
- Jöreskog, K. G. & Sörbom, D. (1981). *LISREL V: Analysis of linear structural relationships by maximum likelihood and least squares methods* (Research Report 81–8). Uppsala, Sweden: University of Uppsala, Department of Statistics.
- Jöreskog, K. G. & Sörbom, D. (1989). *LISREL 7 user's reference guide*. Chicago: SPSS Publications.
- Juel, C. (1991). Beginning reading. In R. Barr, M. L. Kamil, P. B. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research* (pp. 759-788). New York: Longman.
- Kagan, D.M. (1992). Implications of research on teacher beliefs. *Educational Psychologist*, 27(1), 65-90.
- Kaplan, D. (2000). *Structural equation modeling: Foundation and extensions*. Thousand

Oaks, CA: Sage Publications.

- Klingner, J. K., Arguelles, M.E., Hughes, M.T., & Vaughn, S. (2001). Examining the schoolwide “spread” of research based practices. *Learning Disability Quarterly*, 24(1), 221–234.
- Krueger, A. (2000). *Economic considerations and class size*. Princeton University, Industrial Relations Section, Working Paper #447.
- Krueger, A. & Hanushek, E. (2000). *The class size policy debate: understanding the magnitude and effect of class size on student achievement*. Economic Policy Institute, Washington, DC.
- Krueger, A. & Whitmore, D. (2001, January). The effect of attending a small class in the early grades on college test-taking and middle school test results: Evidence from Project STAR. *Economic Journal*, 111, 1-28.
- Kuperminc, G., Leadbeater, B. & Blatt, S. (2001). School social climate and individual differences in vulnerability to psychopathology among middle school students. *Journal of School Psychology*, 39(2), 141-159.
- Kuperminc, G., Leadbeater, B., Emmons, C. & Blatt, S. (1997). Perceived school climate and difficulties in the social adjustment of middle school students. *Applied Developmental Science*, 1(2), 76-88.
- Lambert, L. (1998). *Building leadership capacity in schools*. Alexandria, VA: Association of Supervision and Curriculum Development.
- Lashway, L. (July 2002). *Developing instructional leaders*. ERIC Digest 160, Clearinghouse on Educational Management, University of Oregon.
- Lee, V. & Loeb, S. (2000). School size in Chicago elementary schools: Effects on teachers’ attitudes and student achievement. *American Educational Research Journal*. 37(1), 3-31.
- Lee, V. & Smith, J. (1993). Effects of school restructuring on the achievement and engagement of middle-grade students. *Sociology of Education*, 66(3):164–187.
- Lee, V. & Smith, J. (1995). Effects of high school restructuring and size on early gains in achievement and engagement. *Sociology of Education*, 68(4):241–270.
- LeBerge, D. & Samuels, S. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.
- Leithwood, K. (1988). *Description and assessment of a program for the certification of principals*. Toronto, ON: Ontario Ministry of Education.

- Leithwood, K. & Jantzi, D. (2000). The effects of transformation leadership on student engagement with school. *Journal of Educational Administration*, 38(2), 112-129.
- Leithwood, K., Louis, K.S. Anderson, S. & Wahlstrom, K. (2004). *Executive summary: How leadership influences student learning*. Minneapolis, MN: University of Minnesota, Center for Applied Research and Educational Improvement.
- Leitner, D. (1994). Do principals affect student outcomes: An organizational perspective. *School Effectiveness and School Improvement*, 5(3), 219-238.
- Levine, D.U. & Lezotte, L.W. (1990). *Unusually effective schools: A review and analysis of research and practice*. Madison, WI: National Center for Effective Schools Research and Development. (ERIC Document Reproduction Service No. ED330032)
- Lezotte, L.W. (2001). *Revolutionary and evolutionary: The effective schools movement*. [Online] Available at www.effectiveschools.com/downloads/RevEv.pdf
- Lindsay, P. (1982). The effect of high school size on student participation, satisfaction and attendance. *Educational Evaluation and Policy Analysis*, 4(1), 57-65.
- Louis, K. S. & Miles, M. B. (1992). *Improving the urban high school*. London, UK: Cassell.
- Madden, N.A., Slavin, R.E., Karweit, N.L., Dolan, L.J. and Wasik, B.A. (1993). Success for all: Longitudinal effects of a restructuring program for inner-city elementary schools. *American Educational Research Journal*, 30, 123-148.
- Manning, M. L. & Saddlemire, R. (1996). Developing a sense of community in secondary schools. (National Association of Secondary School Principals). *NASSP Bulletin*, 80(584), 41-48.
- Marcoulides, G. A. & Heck, R. H. (1993). Organizational culture and performance: Proposing and testing a model. *Organizational Science*, 4(2), 209-225.
- Marshall, C. & Greenfield, W.D. (1985). The socialization of the assistant principal: Implications for school leadership. *Education and Urban Society*, 18(1), 3-6.
- McCallum, B. (1999). Literacy in four effective schools. *School Leadership & Management*, (19)1, 7-24.
- McEvoy, A. & Welker, R. (2000). Antisocial behavior, academic failure and school climate: A critical review. *Journal of Emotional and Behavioral Disorders*, 8(3), 130- 140.

- Mendro, R. L., Webster, W. J., Bembry, K. L. and Orsak, T. H. (1995). *An application of hierarchical linear modeling to determining school effectiveness*. Paper presented to the Annual Meeting of the American Educational Research Association. San Francisco, CA.
- Meyer, R. H. (1997). Value-added indicators of school performance: A primer. *Economics of Education Review*, 16(3), 283-301.
- Monk, D. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13(2):125–145.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D. & Ecob, R. (1988). *School matters: The junior years*. Somerset, Open Books (reprinted in 1995 by Paul Chapman, London).
- Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The future of children*, 5(2), 113-127.
- Murphy, J. & Louis, K. S. (1994). *Reshaping the principalship: Insights from transformational reform efforts*. Thousand Oaks, CA, Corwen Press.
- Murphy, J. (1990). Principal instructional leadership. In Thurston, P & Lotto, L (eds), *Advances in Educational Leadership*, 163–200, Greenwich, CT, JAI Press.
- National Association of Elementary School Principals. (2001). *Leading learning communities: Standards for what principals should know and be able to do*. Alexandria, Virginia.
- National Center for Education Statistics. (May 2000). *School-level correlates of academic achievement*. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education. (NCES 2000-303).
- National Commission on Education (1995). *Success against the odds*. London, UK: Routledge and Kegan.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.
- No Child Left Behind Act of 2001*, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
- Oxley, D. (1994). Organizing schools into small units: Alternatives to homogeneous grouping. *Phi Delta Kappan*, 75(7), 521–526.

- Peters, T. J. & Waterman, R.H. (1982). *In search of excellence*. New York, NY: Harper and Row.
- Peterson, K. D., Gok, K. & Warren, V. D. (1995, April). *Principals' skills and knowledge for shared decision making. Paper prepared at the Center on Organization and Restructuring of Schools*. University of Wisconsin, Madison. (ERIC Document Reproduction Service No. ED 386 827)
- Peterson, R.L. & Skiba, R. (2001). Creating school climates that prevent school violence. *The Clearing House*, 74(3), 155-163.
- Phillips, G. W. & Adcock, E. P. (1997). *Measuring school effects with hierarchical linear modeling: Data handling and modeling issues*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.
- Pitner, N. J. (1988). The study of administrator effects and effectiveness. In N. J. Boyan (Ed.), *Handbook of research on educational administration* (1st ed., pp. 99-122). New York: Longman.
- Pittman, R. & Haughwout, P. (1987). Influence of high school size on dropout rate. *Educational Evaluation and Policy Analysis*, 9(4), 337-343.
- Purkey, S. & Smith, M. (1983). Effective schools: A review. *The Elementary School Journal*, 83(4), 427-452.
- Raudenbush, S. W. and Bryk, A. S. (2002). *HLM 5: Hierarchical Linear Models: Applications and Data Analysis Methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Raudenbush, S. W. Bryk, A. S., Cheong, Y. F. and Congdon, R (2001). *HLM 5: Hierarchical Linear and Nonlinear Modeling*. Lincolnwood, IL: Scientific Software International.
- Rea, P. J., McLaughlin, V. L. & Walther-Thomas, C. S. (2002). A comparison of outcomes for middle school students with learning disabilities in inclusive and pullout programs. *Exceptional Children*, 68, 203-222.
- Renchler, R. (1992). School leadership and student motivation. ERIC digest 71. (ERIC Clearinghouse on Educational Management, ED346558)
- Reynolds, D. & Stringfield, S. (January 1996). Failure-free schooling is ready for take-off. *Times Educational Supplement*, 19, 10.
- Roderick, M. & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low-achieving students to high-stakes testing. *Educational Evaluation and Policy Analysis*, 23(3), 197-227.

- Rogosa, D. R., Brandt, D. & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin*, 90, 726–748.
- Rothstein, R. (2001). Why spending doesn't improve test scores. *Principal*, 81(1), 62-63.
- Rowan, B., Chiang, F.S. & Miller, R.J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education*, 70, 256–284.
- Rutter, M., Maughan, B., Mortimore, P., Ouston, J. & Smith, A. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. Boston: Harvard University Press.
- Samdal, O., Nutbeam, D., Wold, B. & Kannas, L. (1998). Achieving health and educational goals through schools-a study of the importance of the school climate and the students' satisfaction with school. *Health Education Research*, 13(3), 383-397.
- Sammons, P., Hillman, J. & Mortimore, P. (1995). Key characteristics of effective schools: A review of school effectiveness research. London, England: Office for Standards in Education (ERIC Document Reproduction Service No. ED389926)
- Sammons, P., Thomas, S. & Mortimore, P. (1997). *Forging links: Effective schools and effective departments*. London, UK: Paul Chapman.
- Sanacore, J. (1997). Guidelines for successful reading leaders. *Journal of Adolescent & Adult Literacy*, 41(1), 64–68.
- Sanders, M. G. & Harvey, A. (2002). Beyond the school walls: A case study of principal leadership for school-community collaboration. *Teachers College Record*, 104(7), 1345-1368.
- Sanders, W.L. & Horn, S.P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, 12(3), 247-256.
- Scheerens, J. & Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon Press.
- Scheurich, J. J., Skrla, L. & Johnson, J. F. (2000). Thinking carefully about equity and accountability. *Phi Delta Kappan*, 82(4), 293-299.
- Scheurich, J.J. (1998). Highly successful and loving, public elementary schools populated mainly by low-SES children of color: Core beliefs and cultural characteristics. *Urban Education*, 33(4), 451-491.

- Schultz, T.W. (1961). Investment in human capital. *The American Economic Review* 1(2), 1-17.
- Senge, P. (1990). *The fifth discipline*. New York, NY: Doubleday.
- Sergiovanni, T. (1991). *The principalship: A reflective practice perspective (2nd ed.)*. Boston: Allyn and Bacon.
- Sergiovanni, T. (1999). *Rethinking leadership: A collection of articles*. Arlington Heights, IL: Skylight Professional Development Publishing Company.
- Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford.
- Schumacker, R. E. & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Snow, C., Burns, M. and Griffin, P. (Eds.), (1998). *Preventing reading difficulties in young children*. Washington DC: National Academy Press.
- Snow, C. E., Barnes, W. S., Chandler, J., Goodman, I. F. & Hemphill, L. (1991). *Unfulfilled expectations: Home and school influences on literacy*. Cambridge, MA: Harvard University Press.
- Snyder, K. J., Acker-Hocevar, M. & Snyder, K. M. (1996, Winter). Principals speak out on changing school work cultures. *Journal of Staff Development*, 17(1), 14-19.
- Spillane, J., Halverson, R. and Diamond, B. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher*, 30 (3), 23-28.
- Stallings, J. & Mohlman, G. (1981). *School policy, leadership style, teacher change and student behavior in eight schools*. Final report to the National Institute of Education, Washington D.C.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173-180.
- Stringfield, S. (1995). Attempting to enhance students' learning through innovative programs: the case for schools evolving into high-reliability organizations. *School Effectiveness and School Improvement*, 6(1), 67-96.
- Stronge, J. H. (1988). A position in transition? *Principal*, 67(5), 32-33.
- Taylor, D. L., & Tashakkori, A. (1995). Decision participation and school climate as predictors of job satisfaction and teachers' sense of efficacy. *Journal of*

- Experimental Education*, 63, 217–230.
- Teddlie, C. & Reynolds, D. (2000). *The international handbook of school effectiveness research*. New York, NY: Falmer Press.
- Teddlie, C. & Stringfield, S. (1993). *Schools do make a difference: Lessons learned from a 10-year study of school effects*. New York, NY: Teachers College Press.
- University of Oregon. (2005). *Oral reading fluency and retell fluency*. [Online] Available at <http://dibels.uoregon.edu/measures/orf.php>.
- U.S. Department of Education (n.d.). *2005 Reading First program funding*. Retrieved [Online] Available at <http://www.ed.gov/programs/readingfirst/funding.html>.
- U.S. Department of Education (May 1998). *Reducing class size: What do we know?* [Online] Available at http://www.ed.gov/pubs/ReducingClass/Class_size.html
- Van de Grift, W. & Houtveen, A. A. M. (1999). Educational leadership and pupil achievement in primary education. *School Effectiveness and School Improvement*, 10(4), 373-389.
- Voelkl, K. (1995). School warmth, student participation and achievement. *Journal of Experimental Education*, 63, 127-138.
- Wagstaff, M., Melton, J., Lawless, B. & Combs, L. (1998). *African-American Student Achievement Research Project*. Kilgore, TX: Education Service Center Region 7, (ED 425 256).
- Wang, M.C., Haertel, G.D. & Walberg, H.J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63(3), 249-294.
- Waters, T., Marzano, R.J. & McNulty, B. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement*. Working Paper. Mid-continent Regional Educational Laboratory.
- Weber, G. (1971). *Inner city children can be taught to read: Four successful schools*. Washington, D.C.: Council for Basic Education, Occasional Papers #18.
- Wenglinsky, H. (1997). How money matters: The effect of school district spending on academic achievement. *Sociology of Education*, 70(3), 221-237.
- Willett, J. B. & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, 116, 363–381.
- Witziers, B., Bosker, R.J. & Kruger, M.L. (August 2003). Educational leadership and

student achievement: The elusive search for an association. *Educational Administration Quarterly*, 39(3), 398-425.

Wolfe, M. P., Bartell, C. A. & DeBolt, G. P. (2000). School, district and University cultures and responsibilities (mentoring framework: Dimension II). In S. J. Odell & L. Huling (Eds.), *Quality mentoring for novice teachers* (pp. 89-95). Indianapolis: Kappa Delta Pi.

Yap, K.O. & Enoki, D. (1995). In search of the elusive magic bullet: Parental involvement and student outcomes. *The School Community Journal*, 5(2), 97-106.

BIOGRAPHICAL SKETCH

Steve Nettles is the Director of Research at the Florida Center for Reading Research at Florida State University (FCRR). He has held positions as a classroom teacher and school administrator, as well as working in program monitoring and evaluation at the school district level. As a senior educational consultant, Dr. Nettles has worked with schools, school districts, and state departments of education across the country on matters of educational leadership and policy. His work at FCRR is focused on developing useful data sources and leadership strategies to promote successful school and district leadership in reading. In this effort, Dr. Nettles has directed the development of an elementary classroom reading walkthrough protocol for principals. He has also worked to collect and analyze statewide data on principal behaviors in the implementation of effective reading programs to inform policy, technical assistance and professional development activities in this area. In addition, Dr. Nettles develops training opportunities for principals to support them in the critical areas of reading leadership.