

THE RELATIONSHIP BETWEEN COACHING HOURS AND CHILDREN'S LITERACY
OUTCOMES, TEACHER PRACTICES, AND/OR CHANGES IN THE PRESCHOOL
CLASSROOM ENVIRONMENT

by

ELIZABETH ROCHELLE CLARKSON

(Under the Direction of Sally J. Zepeda)

ABSTRACT

The purpose of this study was to analyze existing data from one quasi-experimental study that was part of an Early Reading First grant to understand the relationship of coaching hours on child outcomes, teacher practices, and/or changes in the classroom environment. Research questions included 1) What are the relationships between coaching hours and children's literacy outcomes? and 2) What is the relationship between coaching hours and teacher practices and/or changes in the preschool classroom environment? Child outcomes were measured by significant increases in the posttest scores of children's Peabody Picture Vocabulary Test, Fourth Edition (PPVT), the Expressive Vocabulary Test, Second Edition (EVT), and the Phonological Awareness and Literacy Screening (PALS Pre-K) assessments. Teacher practices and classroom environments were measured by changes in the classroom Early Language and Literacy Classroom Observation (ELLCO) posttest scores.

Findings suggested that in the area of receptive language, students made significant gains pretest to posttest, but the interaction and group effects were not significant. For expressive oral language, findings indicate statistically significant differences in the time analysis and in the

group effect. On Rhyme Awareness, the time effect between pretest and posttest and the interaction term between Rhyme Awareness and condition were both significant; however, there was not a statistically significant difference for the group effect.

Findings also indicated positive correlations between Observation and the PPVT and EVT posttest scores. However, there were also several statistically significant negative correlations including: Public Practice and Pre and Post Conferencing to the PPVT posttest scores, Pre and Post Conference to the EVT posttest scores, and Pre and Post Conference and Study Groups to the Print and Word Awareness section of the PALS Pre-K. No statistically significant differences in the time analysis, interaction term, or group effect on any of the classroom environment variables were found. On the ELLCO measure, there were no statistically significant correlations between the number of coaching hours and ELLCO posttest scores, including teachers who participated in Years 1, 2, and 3 of the grant.

INDEX WORDS: Early Childhood Education and Literacy Coaching; Literacy Coaching; Literacy Coaching and Time; Time, Literacy Coaching and Student Achievement; Literacy Coaching and Children's Literacy Outcomes.

THE RELATIONSHIP BETWEEN COACHING HOURS AND CHILDREN'S LITERACY
OUTCOMES, TEACHER PRACTICES, AND/OR CHANGES IN THE PRESCHOOL
CLASSROOM ENVIRONMENT

by

ELIZABETH ROCHELLE CLARKSON

Bachelor of Science, Appalachian State University, 2000

Master of Arts, University of Colorado, 2006

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF EDUCATION

ATHENS, GEORGIA

2013

© 2013

Elizabeth R. Clarkson

All Rights Reserved

THE RELATIONSHIP BETWEEN COACHING HOURS AND CHILDREN'S LITERACY
OUTCOMES, TEACHER PRACTICES, AND/OR CHANGES IN THE PRESCHOOL
CLASSROOM ENVIRONMENT

by

ELIZABETH ROCHELLE CLARKSON

Major Professor: Sally J. Zepeda

Committee: Stacey Neuharth-Pritchett
Catherine C. Sielke

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2013

DEDICATION

The Shirelles may have sung, “This is dedicated to the one I love,” but they don’t know true dedication until they meet Bob and Jackie Clarkson. My dad, a man of few words, taught me to “K.I.S.S.” my school work and my mom, a woman of great force and strength, taught me to fight through what was most difficult in life. Without their quiet support and thundering encouragement, I would not have had the courage to walk into Hudson Elementary School and introduce myself as Elizabeth Clarkson. The list of inspiring driveway pep talks and gestures of embarrassing pride date back to my childhood. From the monotonous athletic events, the race to the top of the hill to hold the signs, the late night phone calls to keep me awake on the long car rides home, to the phone read-aloud-editing, my parents stand as a shining example of love. It is through their dedication to me that I stand before them humbled and grateful for the passion of learning they kindled in me many years ago. For the gifts of persistence and stubbornness, I am forever grateful. Thank you.

ACKNOWLEDGEMENTS

There is a bit of poetic irony that in the section I've been writing in my head for three years, words now fail me. To simply "acknowledge" the people in my life for their support and friendship seems small and fragile compared to the fierce loyalty I feel for them. As the expression goes, actions speak louder than words. I could scream these words of gratitude into the wind, but the expression holds true; your acts of support have been deeply felt and intensely appreciated.

I once stayed up too late talking with a man I found oddly fascinating. I'd met the best friend I'd ever know. Jamie, to me, you are a man of legends. No request too small or too large, you kept up with the mundane tasks to keep the house running, listened to hours of a tired student's rambling on about interesting theories, adding your own insightful ideas, debated the merits of a word change or comma placement, walked the dog, and cooked dinner. Amazing. You gave me the gift of sanity, and I owe you my imagination.

In January of 2010, I bought a house but inherited my Pinebreeze Hood. Our little cul-de-sac continues to bring me smiles from surprise letters in my mailbox and the occasion stray toy on my front porch. The list of support is a long one: the hamburger sliders and freezer drinks you brought me late on summer nights when you saw my light still on, the highlighters you left in my mailbox as quiet reminders that you were cheering for me, the hugs your children threw at me and excitedly asked, "how many more pages do you have?" and the text pictures you sent of a warm and safe pup when you "Derby-sat" for me when I traveled overnight. These acts

of kindness are ones forever etched in my heart. I could never have asked for such support, but you gave it willingly. I am honored to call you all my friends.

Sometimes the stars align in ways that seem impossible and on dark nights after class in Athens, I knew a warm house, soft couch, and shouts of joy awaited me at the house on Crestwood Drive. You generously opened your home and your family to me. My wary mind was rewarded by the intellectual stimulus our conversations offered and my tired soul was comforted by your acceptance. I have only a lifetime of friendship to offer in return.

I feel unapologetically guilty by the quality of instruction offered by my professors at the University of Georgia. They gently and consistently provided me with an academic identity that I could try on and safely test out, but in one that I now firmly stand. Dr. Jack Parish listened to a hopeful, prospective student's bureaucratic hurdles and one at a time, cut away the red tape that made this journey possible. Thank you to my committee, Dr. Zepeda, Dr. Neuharth-Pritchett, and Dr. Sielke for their patience, responses to my endless questions and emails, and encouragement. Dr. Neuharth-Pritchett offered content expertise and methodological insight rarely found as such a powerful combination in the field of early education and without her strong presence this study would not be possible. The pearls of wisdom offered by my Major Professor, Dr. Sally J. Zepeda, could adorn the grandest of queens, and I will be forever awed at her single-minded dedication to her students. I hope we can continue to work together for many years, and I proudly add my name, as number 93, to the long list of Z-Pack members.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER	
1 INTRODUCTION	1
Introduction.....	1
Statement of the Problem.....	3
Purpose of the Study	5
Background of the Study	5
Research Questions.....	9
Significance of the Study	9
Definition of Terms.....	10
Limitations of the Study.....	14
Overview of the Research Procedures	15
Organization of the Dissertation	16
2 REVIEW OF RELATED LITERATURE.....	17
Introduction.....	17
Overview of Early Reading First	18
Coaching as a Form of Professional Development.....	20

	Assumptions of Professional Development and Student Achievement	24
	Review of Quantitative Studies	26
	Review of Studies Specific to the Field of Early Childhood Education.....	33
3	METHODOLOGY	39
	Research Questions	39
	Research Design.....	39
	Early Reading First	40
	Context Specific Early Reading First	42
	Participants.....	44
	Measures	50
	Data Collection	58
	Data Analysis	59
4	FINDINGS.....	60
	Introduction.....	60
	Results from the Longitudinal Study of Early Reading First (GRANT).....	60
	Results from Secondary Data Analysis.....	62
	Summary	75
5	DISCUSSION OF THE FINDINGS	77
	Summary of the Findings	78
	Major Findings Related to Key Literature and Discussion of the Findings	81
	Limitations to the Current Study.....	85
	Implications and Recommendations for Researchers	85
	Implications and Recommendations for Practitioners	87

Implications and Recommendations for Policymakers.....	88
Final Thoughts	89
REFERENCES	91

LIST OF TABLES

	Page
Table 1: Demographic Information for All Teachers	47
Table 2: FY 2011-2012 GRANT Classrooms	49-50
Table 3: Scores and/or Subscales for the PALS Pre-K	55
Table 4: Results of Repeated Measures Analysis of Variance for the PPVT, EVT, and PALS Pre-K Measures	65-66
Table 5: Correlation Analysis for Coaching Practices and for the PPVT, EVT, and PALS Pre-K Measures	68
Table 6: Results of Repeated Measures Analysis of Variance for the ELLCO Measure	71
Table 7: Correlation Analysis for Coaching Practices and for the ELLCO Measure	72-73
Table 8: Results of Repeated Measures Analysis of Variance for the ELLCO Measure Over Time	74

LIST OF FIGURES

	Page
Figure 1: Subcategories Listed in the Coaching Log.....	12
Figure 2: Theory of Action for Professional Development	25
Figure 3: Local Non-Profit Coaching Model.....	44
Figure 4: Graphic Representation of the ELLCO Measure Over Time	75

CHAPTER 1

INTRODUCTION

Introduction

There is no question that the most recent iteration of the Elementary and Secondary Education Act of 1965 (ESEA), the federal law known as No Child Left Behind (NCLB) (2001), changed the landscape of American education not only in terms of federal involvement, high expectations for teachers and students, but also for fiscal investments in education (DeBray-Pelot & McGuinn, 2009; DellaMattera, 2010). The NCLB Act responded to the nation's dilemma of school readiness and the accountability movement spurred by this Act filtered down to influence early education, including the preschool and early learning space through the requirement that all states create early education standards or preschool policy guidelines. DellaMattera (2010) found in her study of New England states' preschool policy guidelines that there is a "significant emphasis placed on children's cognitive development" as evidenced through the language in their guidelines (p. 38). Thus, it is not surprising that NCLB "thrust early education squarely into the political arena" (DellaMattera, 2010, p. 46).

One example of a discretionary federal grant born from NCLB in 2002 was Early Reading First (ERF). ERF was part of the President's "Good Start, Grow Smart" initiative and was "designed to transform existing early education programs into centers of excellence that provide high-quality, early education to young children, especially those from low-income families" (U.S. Department of Education, 2012, ¶ 3). Through an extensive and highly competitive grant process, federal funds were awarded to local programs that showed they had

the strategic plan and capacity to “enhance young children's language and cognitive development by providing high-quality instruction and ongoing professional development based on scientifically based research” (U.S. Department of Education, 2012, ¶ 2).

Shidler (2009) cited that “increased teacher efficacy (both instructional and self) has been an essential component to various educational reforms” (p. 453) including legislation such as No Child Left Behind and federal grants such as Reading First projects and Early Reading First Projects. The use of coaching as a component of professional development has also been employed as a means to improve teacher performance and student outcomes on various assessments (Shidler, 2009). According to the Final Report to Congress in 2007, there were two key elements of Early Reading First—the use of scientifically based methods and the goal of enhanced professional development. ERF included not only the expectation for professional development but also that professional development for teachers was expected to be “continuous, intensive, and classroom focused” (Jackson, McCoy et al., 2007, p. xiii).

One grant recipient, a local non-profit, was awarded a total of four Early Reading First grants, beginning in 2005. For the purposes of this research, the local non-profit was referred to as LNP. Early Reading First is uniquely situated between the development of policy and the implementation of practice. While the federal government provides overall goals and boundaries, individual fiscal agencies are left to decide on and implement the most effective strategies on how to use their resources and dollars to accomplish these goals. Part of LNP’s approach to the work of professional development in their ERF grants included providing coaches who would work in classrooms and four days of early literacy training for teachers to help them fully integrate literacy strategies throughout their classrooms.

Early childhood programs, especially programs serving Pre-Kindergarten children, are routinely expected and called on to improve significantly the literacy skills and general school outcomes for children of low-income status and at risk for academic challenges (Powell, Diamond, Burchinal, & Koehler, 2010). The assumption of literacy coaching embraces that “through the job-embedded, ongoing professional development provided by literacy coaches, classroom teachers will improve their instruction, which will lead to increased student achievement” (Elish-Piper & L’Allier, 2011, p. 84). Through this basic assumption of coaching intervention, it is hoped that teachers will yield improvements in their instructional capabilities that will eventually increase gain scores and improve student achievement; these premises have been investigated in a small, but growing number of studies (Powell et al., 2010).

Statement of the Problem

More than two decades ago, Carpenter, Fennema, Peterson, Chiang, and Loef (1989) produced a “groundbreaking study” demonstrating that professional development could improve student achievement (cited in Wayne, Yoon, Zhu, Cronen, & Garet, 2008, p. 469). If coaching continues to be used as a professional development strategy, more detailed and specific research is critical to responsibly using that capital as effectively as possible. Although coaching, usually embedded in mentoring programs, has become more accepted in early childhood settings, there is a lack of experimental studies focusing on coach-teacher processes and outcomes (Landry, Anthony, Swank, & Monseque-Bailey, 2009). Wayne et al. (2008) agreed that there is a lack of research to support and/or provide guidance as to the most high yield strategies in coaching because coaching as a strategy for improved student outcomes is still inconclusive at this time. Wayne et al. (2008) explained, “Given the large public investment in professional development,

there is much to gain from research that addresses practical questions faced by those who design and adopt professional development programs” (Wayne et al., 2008, p. 476).

Specific to early childhood, Susan Neuman, former Assistant Secretary of Elementary and Secondary Education under the Bush administration and author of *Early Reading First*, explained that there is ample research on the process of literacy coaching but limited evidence to its effectiveness in early childhood settings (Neuman & Cunningham, 2009). The Director of the National Institute for Education Research, Dr. Steve Barnett, cited that large scale programs have often failed to bring to scale the types of programs necessary to deliver promised returns on investment (Barnett, 2012). He explained that often times programs did not duplicate the models that proved successful in research because these failed programs focused on cheaper implementation, sacrificing proper design and evaluation.

Several questions piqued the researcher’s interest in Early Reading First grants related to coaching and literacy. Early Reading First grants are only one of many public programs, but did this grant achieve higher student outcomes as a result of increased professional development through their use of a coaching model? To address a more specific component, what was the relationship between coaching hours and student achievement, teacher practices, and/or changes in the classroom environment? This last question, while simply posed, abounds with complex webs of independent and dependable variables. Although aspirational, the cumulative body of research could allow for more effective coaching practices to surface, to be replicated, and to be brought to scale when appropriate. More effective forms of professional development opportunities for teachers, specifically within a coaching model, might have the capacity to produce the overarching goal of increased student achievement.

Purpose of the Study

The purpose of this study was to examine one Early Reading First project and to analyze the available data to shed light on the relationship of coaching hours and student achievement, teacher practices, and/or changes in the classroom environment. The findings of such a study could contribute to the knowledge gap in the field of early childhood by isolating specific variables of professional development, including coaching hours, and analyzing them against posttest scores within an intervention and comparison group of classrooms.

Acknowledging the importance of a quality early childhood school experience, there is serious misalignment in the education of some early childhood educators and the necessary skills to “optimize classroom practice” (Landry et al., 2009, p. 449). Professional development, as Landry et al. (2009) explained, may “serve as the buffer” and offer teachers the necessary skills to deliver effective instruction and to prepare young students for school success (p. 449). Coaching may serve as an effective form of professional development that helps bridge the gap between what Landry et al. (2009) called “inadequate teacher preparation” in the field of early childhood education (p. 449).

In the current study, coaching is referenced as a form of professional learning embedded in practical educational applications and settings designed to increase the capacities of teachers to use curricula and teaching strategies to improve student achievement. This study focused on the professional development strategy of coaching as a means to increase student achievement, as evidenced by secondary data analysis of one Early Reading First grant.

Background of the Study

The desire to increase student achievement in school is not new to education, nor is the idea of accomplishing this task through systems of professional development. Joyce and

Showers (2002) proposed that when teachers engaged in professional development around a curriculum area or teaching strategy that was useful across the curriculum areas and regularly studied the implementation and resulting student learning, the “odds of student achievement will rise substantially” (p. 3). These foundational authors made clear that professional development and student achievement were vitally and causally linked together. Educators have the knowledge to design and to implement successful programs that will result in improved educational experiences for students so a richer body of research will add support to their claims.

Beginning in the 1980s, Joyce and Showers (2002), through an exhaustive literature review of professional development training, presented the hypothesis that training components with “presentation by theory, modeling or demonstration, practice, structured and open ended feedback, and in-class assistance with transfer” were the types of training most likely to lead to various levels of impact (p. 85). If teachers were expected to master new curricula, approaches of teaching, and new practices, then technical assistance at the classroom level would be critical.

Joyce and Showers’ research in the early 1980s confirmed that teachers with technical assistance in the classroom are more likely to exhibit greater implementation than those who shared initial training but did not have the support of a coach. Shanklin (2009) explained that coaching emerged as a “sensible means to increase teacher quality and subsequent student learning (p. 42).

Coaching, professional development, and teacher effectiveness are well documented to be intertwined and linked with student achievement (Elish-Piper & L’Allier, 2010; Joyce & Showers, 2002; Kretlow & Bartholomew, 2010; Shidler, 2009; Vanderburg & Stephens, 2010). The strategy of coaching, as a form of professional development designed to improve teacher practices and student achievement, has been widely adopted as a process to support both teacher

and student learning (Neuman & Wright, 2010; Zepeda, 2012a). Despite the positive emotion, studies of the impact of coaching suggested mixed results in both early childhood settings and elementary grades, Kindergarten through third grades.

As recently as 2011, Elish-Piper and L'Allier examined the available literature that focused on the relationship between literacy coaching and teacher practices. Specifically, Elish-Piper and L'Allier's study focused on the relationship between the amount, time, and content of literacy coaching and Kindergarten through third grade reading gains. Based on their review, coaching hours may be positively related to the enhancement of classroom literacy environments, the use of best practice literacy instruction, and the implementation of new instructional practices. This study and others suggested a strong relationship between the coaching strategy and teacher professional development and expertise.

While Elish-Piper and L'Allier (2011) stated that none of these studies specifically examined the impact of literacy coaching on student achievement, they did report that certain aspects of literacy coaching predicted student reading gains in more than one grade between Kindergarten and third grade. Bean, Draper, Hall, Vandermolen, and Zigmond (2010) found a significant relationship between the amount of coaching performed in schools and student achievement, but they caution there are still unanswered questions about to what extent the two events are actually related.

In contrast, several studies reported little to no significant correlations between coaching hours in the classroom and student achievement (Landry, Swank, Smith, Assel, Gunnewig, 2006; Neuman & Cunningham, 2009; Neuman & Wright, 2010; Shidler, 2009). In year two of Shidler's (2009) three year preschool research study, she reported no statistically significant correlation between gain scores and hours spent coaching per classroom. Overall, Shidler

concluded that "a more focused, honed approach to coaching teachers in enhancing child outcomes in specific measures was more effective" (Shidler, 2009, p. 459).

Despite the differing research conclusions, Landry et al. (2006) agreed with the powerful potential of coaching, citing that the National Association for the Education of Young Children (NAEYC) and the International Reading Association (IRA) noted that coaching and mentoring focused on a curriculum, instruction, and assessment tools, and relationships are a "necessary part of a complete professional development package" (p. 308). Neuman and Cunningham (2009) also recognized coaching as an essential component of any effective professional development intervention. The authors emphasized that if we, as educators, seek to improve children's skills in school, we need to be certain teachers have the foundational skills necessary to provide high quality instruction. Neuman and Wright (2010) offered a warning, "We need to be cautious; therefore, when extrapolating from observed changes in the environment to likely changes in children's performance" (p. 84).

Deussen, Coskie, Robinson, and Autio (2007) explained that coaching qualifications, responsibilities, and time allocation of literacy coaches vary widely between settings, and those differences "are often disregarded in the literature, making it more difficult to interpret findings about implementation success and impact on both teachers and students" (p. 6). While the connectedness of professional development and student achievement is often accepted as truth, little research exists to offer a research perspective on its actual results, specifically the relationship between coaching hours and student achievement, teacher practices, and/or changes in the classroom environment.

Research Questions

The literature indicates that individualized and sustained work that provides teachers with guidance and feedback on implementation of evidence-based practices in their own classrooms is superior to one-time workshops in improving teacher quality (Powell et al., 2010), yet the question remains: To what extent are coaching and student achievement related, given the assumptions previously described?

Intensive and sustained forms of professional development such as coaching with teachers are of particular interest; however, this evidence-based knowledge has been “widely ignored in early childhood programs” (Powell et al., 2010, p. 299). To add to the research and evidenced-base of this practical coaching conversation, the following questions were addressed in this study.

- 1) What are the relationships between coaching hours and children’s literacy outcomes?
- 2) What is the relationship between coaching hours and teacher practices and/or changes in the preschool classroom environment?

Significance of the Study

From the research perspective, the significance of this study lies in its ability to add to the field’s understanding and knowledge about the effect that dosages of different types of coaching activities have on student outcomes. Neuman and Wright (2010) cited the need for future research looking specifically at coaching dosage to be of critical importance. The evidence on necessary specific features of effective professional development is weak and the body of work as a whole falls short in addressing several practical questions (Wayne et al., 2008).

The practical significance of this study drives this line of inquiry in light of the emphasis on school readiness for children entering Kindergarten. Many literacy initiatives include and

often times focus their professional development intervention around coaching. Extensive literature on the topic provided direction and guidance on successful strategies and principles to consider when offering meaningful professional development to teachers, but without the practical application and connection of real-world coaching data and student results, this literature runs the risk of remaining isolated within research journals and unused by teachers and coaches. With practical data-driven results, initiatives currently in progress may consider course corrections in the way they are allocating coaching time, in the types of activities coaches are involved, and in considering the hours spent in relation to student achievement as valuable variables worthy of serious attention.

There are also implications for policy development and initiative design in the field of early childhood education. Specifically, this study's questions, statistical examination, and analysis of findings could add to the knowledge base that provides direction to policy makers and designers of coaching programs for early childhood educators. This data analysis template could also serve as a format for other Early Reading First grants or early literacy initiatives to review their own use of coaching and its achievement relationships within their own contexts. Equally important, the contribution to the field of quantitative early childhood research has the potential to provide additional direction to support future grants and initiatives to take full advantage of this type of federal investment in education reform.

Definition of Terms

GRANT: GRANT is a pseudonym for one of several ERF grant awarded to LNP. GRANT was awarded to LNP in 2009. GRANT finished its third year of implementation in 2012, with data collected from October 2011-May 2012.

Coaching: A form of professional learning embedded in practical educational applications and settings designed to increase the capacities of teachers to use curricula and teaching strategies to improve student achievement.

In the case of LNP, a coach was a person hired and employed by LNP to work with a specific set of teachers within the context of the individual teachers' classrooms.

Coaching Subcategories: The coaching log categorizes in what ways the coaches spent their time with the teachers, for example public practice or observing teachers. Elish-Piper and L'Allier (2011) refer to these specific activities as the *type of activity*. Elish-Piper and L'Allier (2011) studied literacy coaching within an elementary school context, focusing on grades Kindergarten through third grade. In their study, they divided the information from the coaches' comprehensive list of activities into three categories; *type of coaching activity* (modeling, observing), *context* (study group, individual meeting), and *content* (comprehension, writers' workshop). This study drew on Elish-Piper and L'Allier's descriptions, but the emphasis was placed on the type of coaching activity, referred in the current study as subcategories of coaching. While the other two categories of context and content were important, the structure of the ERF coaching log was designed to capture explicitly information pertaining to the type of activities the coaches were engaged. As much as possible, the other two categories of context and content were teased out of the coaching logs, included, and used as appropriate in the analysis.

These activities or subcategories as described in the coaching log were measured in actual hours and minutes and recorded as time spent. Figure 1 outlines the subcategories recorded in the coaching log, provides a brief description, and includes information on how the hours were coded for analysis.

Public Practice	The term used within LNP to include demonstrations or modeling of a teaching strategy or method of instruction.
Pre and Post Conference	LNP developed a “Reflective Teaching Cycle” to guide coaches’ and teachers’ conversation. Pre conference time was designated specifically for the coach and teacher to plan a lesson and thinking through instruction targets and appropriate assessment measures. A planned observation period followed in which the coach observed the lesson previously planned. The post conference time included jointly assessing the teachers’ implementation of the lesson plan, instructional goals, and assessment results. This time was also dedicated to planning next steps and addressing ways the coach can support the teacher in the future.
Coach Observation	Time dedicated for the coaches to observe in teachers’ classrooms, script conversations or pieces of a lesson, and videotaping lessons.
Data Collection	Collection of various data and assessment information, including both formative (progress monitoring) and summative data.
Progress Monitoring	Activities related to supporting teachers during administration and interpretation of the Individual Growth and Development Indicators (IGDIs), activities related to the teachers’ understanding of using progress monitoring data as a tool for tracking progress, as well as their ability to use the Response to Intervention model.
Study Groups	The ERF grant requires that the coaches offer three hours of study group per site. This time is dedicated to addressing any content area (phonological awareness, vocabulary) or technique (dialogic reading, interactive writing) the coach and/or administration felt were most needed. The study group formats varied, including lecture, hands on activities, reading, and reflection activities. For the purposes of analysis, study group prep time was separated from the 3 required hours of study group delivery.
Administrative Meeting/Duties	This time included time the coaches spent with their site’s administrator.
Office Day	LNP decided that one day a week, the coaches should report to their office and use that time for planning, preparing their documentation (coach log, weekly schedule, coach narrative, etc), and any additional administrative work that was necessary. The day was also held for their weekly GRANT team meeting. For the purposes of the current study’s analysis, these hours were not included in the total amount of coaching time.
Professional Development	This time was specific to the professional development that the coaches’ received, such as attendance at a conference or professional development outside their typical staff meetings. If the coaches attended the GRANT teacher professional development courses, that time was also included in this section.
Other	This category was used to capture any other work time, including planning for workshop presentations and conferences.

Figure 1. Subcategories Listed in the Coaching Log. (McCrosky, 2010).

Coaching Model: A coaching model is a logic model or set of assumptions that guide the coaches in their work with teachers. In the case of GRANT, the coaching model included the Pre Conference, Observation, and Post Conference activities of the coaches.

Instructional Time: Instructional time includes subcategories categories from the coaching log listed as 1) Public Practice, 2) Study Groups, 3) Pre and Post Conference, 4) Coach Observation, and 5) Extended Coaching Sessions. Note that only the first four categories were recorded in the coaching log and used in this research.

Non-instructional Time: Non-instructional time includes subcategories from the coaching log listed as 1) Data collection/Entry, 2) Administrative Meeting and Duties, 3) Study Group Preparation, 4) Progress Monitoring, and 5) Other.

Student Measures: Student measures were pre and post assessments administered to children including the Peabody Picture Vocabulary Test, Fourth Edition (PPVT), the Expressive Vocabulary Test, Second Edition (EVT), and the Phonological Awareness and Literacy Screening (PALS Pre-K). These assessment tools helped to identify children who had developed particular skills and who were ready for additional instruction and likewise, identified children who needed more intensive instruction and/or monitoring in specific areas. These assessments are discussed in more detail in Chapter 3.

Teacher Measures: The teacher measure used in GRANT included the *Early Language and Literacy Classroom Observation* tool (ELLCO) to assess the classroom literacy environments and professional development of the teachers. GRANT also used the OWL Fidelity Checklist, teacher knowledge surveys, and individual goals for teachers as measures of teacher knowledge, growth, and improvement. The OWL Fidelity Checklist was used in the ERF longitudinal study and was used to measure the application of the SBRR curriculum. The

ELLCO is discussed in more detail in Chapter 3. For the purposes of this secondary data analysis, only the ELLCO scores were used to measure changes in teacher practices and/or changes in the environment.

Limitations of the Study

GRANT was quasi-experimental in design, and involved the very real and practical scenarios of teaching; therefore, it focused primarily on implementation. For the current study, the data set used consisted of secondary data collected by an outside evaluator. The evaluator followed the guidelines set forth by LNP, established by a formal memorandum of understanding and contract for services. The researcher was granted permission to analyze the data that were collected, but did not have any influence over the original evaluation design. In places, the data set was not comprehensive but additional site demographics were added when available. Rogers, Anderson, Klinger, and Dawber (2006) explained “consequently, it is not possible to manipulate the conditions and context, even if these variables could be manipulated in practice” (p. 760). Rogers et al. (2006) also pointed out from their own secondary analysis that data sets may have problems outside the analyst’s control, including a lack of relevant variables collected, or missing data due to unresponsiveness. The current study experienced similar challenges, and the researcher relied on the longitudinal evaluation to fill in missing data.

Lastly, in GRANT’S original design and in its authentic teaching environment, there were factors the study did not attempt to account for, such as teacher motivation and parental engagement. As noted in GRANT’s final report, the study experienced a higher rate of teacher attrition than desired, but staff turnover is higher than average in early childhood settings for reasons that are not discussed at this time. Generalizability for this study is limited and other

researchers should carefully consider the geographic area, demographics, and other specifics before comparing to their own group of participants in their unique setting.

Overview of the Research Procedures

The purpose of this study was to examine one Early Reading First project and analyze the available data to shed light on the relationship of coaching hours and student achievement, teacher practices, and/or changes in the classroom environment. Without discounting the value of teacher efficacy and categorical information, this research was focused on the available quantitative data archived from one Early Reading First grant situated in a large urban city in Georgia. Analysis of variance tests and correlation analysis were used to look at the relationship and correlation between coaching hours and students' gain scores on the PPVT, the EVT, and the PALS Pre-K. As appropriate, classroom ELLCO scores, teachers' knowledge surveys, and teachers' attendance in the year-long professional development course work were included.

Using analysis of variance tests, the main effects and interaction effect of coaching hours on the condition and group effects were determined. Using the available coaching data from ERF Years 1-3, the data points from individual teachers, and their students' posttest scores in multiple years were organized and plotted. These data were organized in this matter to determine if the numbers of years enrolled in ERF; therefore, more hours of professional development, were related to the increases in individual student's posttest. All of the data were evaluated for statistical significance, but the effect sizes were also described in terms of their practical importance.

Organization of the Dissertation

Chapter 1 provides the foundation for the organization for this dissertation. The background and statement of the problem anchor the purpose and significance of the study. The introduction also includes key definitions of terms and an overview of the quantitative research procedures used to search for a relationship of coaching hours and student achievement, teacher practices, and/or changes in the classroom environment.

Chapter 2 offers an extensive literature review of professional learning, emphasizing coaching as one strategy or form of job-embedded learning, and the seminal research that has been conducted in early childhood settings, as well as references to research in grades Kindergarten through third grade. Chapter 3 explains the methodology chosen for this research and the basis for the analysis that was described in Chapter 4.

Chapter 5 begins with a discussion of the research findings, including the situated information specific to GRANT and general observations and consequences of the data. Chapter 5 ends with implications for future researchers, practitioners, and policymakers.

CHAPTER 2

REVIEW OF RELATED LITERATURE

Introduction

This study sought to examine one Early Reading First project and analyzed the available data to shed light on the relationship of coaching hours and student achievement, teacher practices, and/or changes in the classroom environment. The secondary analysis also added to the available research and literature base on this relationship. This study and the information gleaned from its results will hopefully add to the knowledge base that Landry et al. (2006) called for to “determine[ing] the relative effectiveness of professional development programs with different durations or different allocations of professional development events across time” (470). The research questions that guided this quantitative study were:

- 1) What are the relationships between coaching hours and children’s literacy outcomes?
- 2) What is the relationship between coaching hours and teacher practices and/or changes in the preschool classroom environment?

Both politicians and educators looking to “advance teacher performance for the purpose of increased student achievement “have included a coaching strategy into their professional development recommendations for grants (Shidler, 2009, p. 453). The increased attention on professional development and higher standards for student achievement are significant reasons literacy coaching surged in popularity.

Powell et al. (2010) asserted that sustained and intense forms of professional development such as coaching were related to improving teacher quality. Substantial literacy research is available on effective professional development characteristics, as well as evidence-based knowledge from the field that link effective teaching practices to student literacy outcomes. However, this knowledge has been “widely ignored in early childhood programs” (Powell et al., 2010, p. 299).

This study acknowledges and appreciates the importance of qualitative research in understanding relationships between two variables, especially in light of the human element, motivations, and passions that surround educating our youngest children. A quantitative analysis for this particular study of an Early Reading First program was chosen to examine the statistical relationships and impact between the variables of coaching, student achievement, teacher practices, and/or changes in the classroom environment over time.

Overview of Early Reading First

One example of a discretionary federal grant born from No Child Left Behind (NCLB) Act in 2002 was Early Reading First (ERF). ERF was part of President Bush’s “Good Start, Grow Smart” initiative and was “designed to transform existing early education programs into centers of excellence that provide high-quality, early education to young children, especially those from low-income families” (U.S. Department of Education, 2012, ¶ 3). Through an extensive and highly competitive grant process, federal funds were awarded through the Office of Early Learning (OEL) to local programs that showed they had the strategic plan and capacity to “enhance young children's language and cognitive development by providing high-quality instruction and ongoing professional development based on scientifically based research” (U.S.

Department of Education, 2012, ¶ 5). According to the U.S. Department of Education, the program goals for ERF are to:

- Support local efforts to enhance the early language, literacy, and pre-reading development of preschool-age children, particularly those from low-income families, through strategies and professional development that are based on scientifically based reading research.
- Provide preschool-age children with cognitive learning opportunities in high-quality language and literature-rich environments so that the children can attain the fundamental knowledge and skills necessary for optimal reading development in Kindergarten and beyond.
- Demonstrate language and literacy activities based on scientifically based reading research that support the age-appropriate development of
 - Oral language (vocabulary, expressive language, listening comprehension)
 - Phonological awareness (rhyming, blending, segmenting)
 - Print awareness
 - Alphabetic knowledge
- Use screening assessments to effectively identify preschool-age children who may be at risk for reading failure (U.S. Department of Education, 2012, ¶ 6).

The National Evaluation of Early Reading First in 2007 suggested that professional development be “continuous, intensive, and classroom focused” (Jackson et al., 2007, p. 27). Although a specific coaching model for Early Reading First is not prescribed to grantees, professional development that included coaching as a strategy was encouraged. The U.S. Department of Education (2003) also listed coaching as an example of professional development

based on scientifically-based reading research. Examples of coaching, as defined in the document outlining guidance to Early Reading First, included “demonstration by the coach of effective strategies, and coaches’ observation of teachers’ instruction followed by discussion and reflection on the effectiveness of instructional strategies and how they support student progress” (U.S. Department of Education, 2003, ¶ 3).

Coaching as a Form of Professional Development

Beginning in the 1980s, Joyce and Showers (2002), through an exhaustive literature review of professional development training, presented the hypothesis that training components with “presentation by theory, modeling or demonstration, practice, structured and open ended feedback, and in-class assistance with transfer” were the types of training most likely to lead to various levels of impact (p. 85). If teachers were expected to master new curricula, approaches of teaching, and new practices then technical assistance at the classroom level would be critical. Their research in the early 1980s confirmed that teachers with technical assistance in the classroom were more likely to exhibit greater implementation than those who shared initial training but did not have the support of a coach.

In addition to their work in professional development, Joyce and Showers (2002) also mentioned other forms of coaching in their research, including technical coaching, challenge coaching, and cognitive coaching, and concluded that all these forms of coaching were similar in that they wished to convey a separation from evaluation. However, all these coaching forms share a feedback component and/or procedures to improve or change classroom practice. The role of evaluation within coaching, even as a perceived notion, exists in intense debate today and heavily influences the field of supervision and instruction (Holland, 2005; McGreal, 1997; Nolan, 1997; Zepeda, 2012a, 2012b).

It is interesting to note that Joyce and Showers (2002) decided to omit feedback as a coaching component and focused on the “collaborative planning and development of curriculum and instruction in pursuit of their shared goals” (p. 88). As a result of this shift in thinking, the authors also redefined the word “coach” to explain that, “when teachers observe each other, the one teaching is the “coach” and the one observing is being “coached”” (Joyce & Showers, 2002, p. 89).

Since this early research and practical applications of coaching, many others have joined the conversation and while debate still continues around the specific details of coaching qualifications, critical components of successful models, and the need for and use of feedback, most people agree with Joyce and Showers that embedding instructional support in the classroom leads to more effective teaching practices and improved child outcomes. According to Shanklin (2009), the Literacy Coaching Clearinghouse (LCC) Advisory Board decided to use the verb form of “coaching” rather than the noun for “coach” to place emphasis on the act of coaching rather than the position. From the early childhood literature related to coaching, there are multiple definitions and variations of the act of coaching. Several examples include:

- a strategy that has the capacity to meet teachers’ needs as it involves modeling, observation, and constructive feedback (Shanklin, 2006)
- a process that involves pairing teachers with a skilled professional in the field who can assist them in developing instructional knowledge and skills (Neuman & Cunningham, 2009)
- “a site-based PD [professional development] initiative designed to develop theory and use demonstration, observation, and feedback to improve classroom practice” (Walpole, McKenna, Uribe-Zarain, & Lamitina, 2010, p. 118)

- “intense form of PD [professional development] (Powell et al., 2010, p. 300)
- “engaged small-group initial training, followed by multiple observations, feedback, and modeling (Kretlow & Bartholomew, 2010, p, 279)
- “master teachers who provide essential leadership for the school’s overall literacy program” (Sturtevant, 2003, p.1)
- “emphasized as an effective way of providing teachers with support and on the job guidance as they engage in their daily teaching activities” (Onchwari & Keengwe, 2008, p.19)
- “a relationship-based process led by an expert with specialized and adult learning knowledge and skills, who often serves in a different professional role than the recipient(s). Coaching is designed to build capacity for specific professional dispositions, skills, and behaviors and is focused on goal-setting and achievement for an individual or group” (National Association for the Education of Young Children (NAEYC) & National Association of Child Care Resource & Referral Agencies (NACCRRA), 2011, p. 11)

For the purposes of this study, the researcher proposed a definition of coaching for this particular context that reflects the statements previously mentioned: a form of professional learning embedded in practical educational applications and settings designed to increase the capacities of teachers to use curricula and teaching strategies to improve child outcomes.

Smith, Robbins, Schneider, Kreader, and Ong (2012) conducted an interview study from technical assistance providers in 17 states to understand the features of quality assistance as part of Quality Rated Improvement Systems. Through their recent review of studies, Smith et al. (2012) stated that while research suggests positive benefits of coaching, there currently lacks “definitive evidence of specific coaching models or features that produce good outcomes” (p. 5).

Part of the shortfall and limitations in current studies stemmed from the lack of detail about coaching methods.

Brown, Stroh, Fouts, and Baker (2005) explained that coaching fell into two main areas, coaching as a component of professional development and coaching aimed at whole school reform. The idea of capacity building fits with the notice of professional development in schools and has been an important element in any professional development strategy or overall plan. As a result of various approaches, functions, organizations, and purposes of coaching in different contexts, coaching as a strategy lacks an exact definition, but has been heavily influenced by the constructivist approaches to learning and teaching (Brown et al., 2005).

Overall, the literature reveals “relatively weak theoretical foundations to the practices [of coaching] and often little attention to the ideas underlying the whole process of coaching” (Brown et al., 2005, p. 8). Often, the task of a researcher included identifying a theoretical model by drawing on the implied or inferred coaching practices. According to the research of Brown et al. (2005), this led to an “eclectic or “common sense” approach based on an individual’s own experiences in a given setting” (p. 8).

The purpose of the discussion on the lack of and/or limited use of a coaching model in this review of the literature is not intended to generate a list of theoretical frameworks used to guide coaching models. The discussion is meant to provide a context for understanding the challenges involved in identifying a theoretical model of coaching used in a practical and, in many times, scaled program setting. The example in this study is the practical and day-to-day activities involved in coaching in early childhood classrooms based on a federal initiative that left the implementation of professional development to the individuals delivering the services. With this open-ended professional development component of ERF in mind, one should

acknowledge the research on coaching, but also consider the underlying assumptions in the relationship between professional development, specifically coaching hours, child outcomes, teacher practices, and/or changes in the classroom environment. The following two sections examine these assumptions of causality and the specific quantitative research in early childhood education that support or call into question those assumptions.

Assumptions of Professional Development and Student Achievement

It is well documented that coaching, professional development, and teacher effectiveness are intertwined and all linked with student achievement (Elish-Piper & L'Allier 2010; Joyce & Showers, 2002; Kretlow & Bartholomew, 2010; Shidler, 2009; & Vanderburg & Stephens, 2010). It is also generally accepted that “intensive, sustained, job-embedded professional development” focused on the content is more likely to improve teacher knowledge, classroom instruction, and student achievement (Wayne et al., 2008, p. 470). A logical extension of this line of thinking would propose that through this professional development and improved teaching practice, students would benefit and demonstrate improved academic success. However, student outcome falls farther down the “chain of linked events” than does changing a teacher’s practice (Landry et al., 2009, p. 463). The components previously listed may seem causally related; the assumption holds that professional development improves teacher practices, which, in turn, improves student outcomes. Figure 2, offered by Garet, Cronen, Eaton, Kurki, Ludwig, Jones, ... and Silverberg (2008) provides a visual representation of this commonly held assumption and its theory of action to illustrate the relationship between professional development and student achievement, teacher knowledge, and teacher practices.

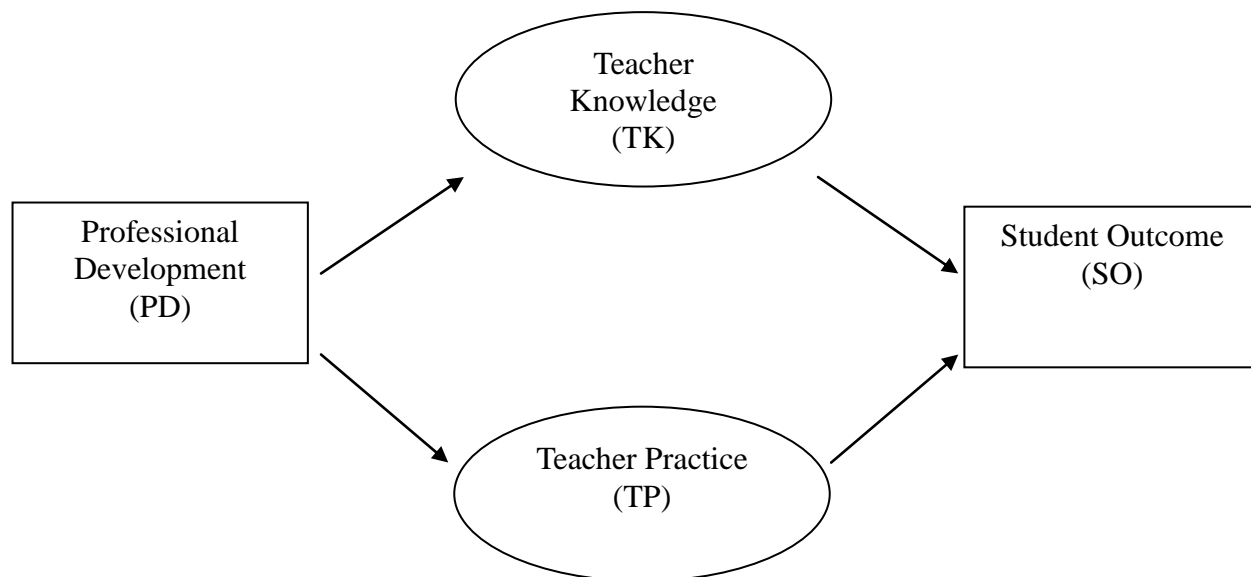


Figure 2. Theory of Action for Professional Development. (Garet et al., 2008).

This relationship between professional development and student outcome may seem to be a logical line of thinking and has been accepted and practiced for more than a decade, but Wayne et al. (2008) warned that this assumption “lacks sufficient specificity to guide practice” (p. 470). As Landry et al. explained, these events are linked, but to what degree is coaching directly correlated with student achievement? The U.S. Department of Education (DOE) and The Institute of Education Sciences (IES) explained that while coaching is an increasingly common approach in professional development, “little is known about its effectiveness” or the specific factors that are involved (Garet et al., 2008, p. 73). They continued by stating that there is “little strong evidence to guide practitioners or researchers” in developing professional development systems that have the potential for promising interventions (Garet et al., 2008, p. 2).

Review of Quantitative Studies

The introduction by No Child Left Behind in 2001 leads to an “intensified spotlight on teacher preparation, reading instruction, and student achievement” and has played a monumental role in bringing the concept of job-embedded professional development via coaches to the national stage (Elish-Piper & L’Allier, 2010, p. 162). The coaching in primary grades is more prevalent than research in early childhood settings and includes the relationships that coaching and professional development have to student achievement or student outcomes. As a result, several studies have emerged examining the relationship between coaching and student achievement (Bean et al., 2010; Elish-Piper & L’Allier, 2011; Elish-Piper & L’Allier, 2010; Walpole et al., 2010). While the concept of coaching enjoys popularity as a strong component of an effective strategy in professional development, studies showed various conclusions and mixed results in the relationships between coaching and student achievement.

A brief review of the literature in middle school and primary grades are useful to the discussion of early childhood research as it provides a history of types of research and supports a foundation for quantitative types of analysis that address coaching as one component of professional development and its impact on student achievement. Two recent studies that examined coaching at the middle and elementary level included Bean et al. (2010) and Lockwood, McCombs, and March (2010). Both of these studies included large sample sizes in either a state or federally funded initiative with positive results on the impact of coaching on student achievement. Both researchers also offered caution in the use of coaching and suggested future considerations for additional research in the coaching arena.

Spurred by the initiative, “Just Read, Florida!” in September 2001, the state experienced an influx of funds designated for literacy coaches to be placed at the lowest performing elementary and secondary levels (Lockwood et al., 2010). The initiative’s goal was for all students to read at or above grade level by 2012. A key component of this initiative was the addition of literacy coaches. Thus, Florida was provided a unique opportunity to “study the effects of coaching on student achievement” (Lockwood et al., 2010, p. 375).

Data in Lockwood et al.’s analysis came from four sources; state literacy coach logs (collected bi-weekly), student achievement data, and two sources of school characteristic data. After merging all data collected, the sample size represented 987 schools in Florida, 644 who had coaches during at least one year during the analysis period. The dataset included 90% of schools containing at least one middle school grade (sixth through eighth grade), and nearly all middle schools that received a state funded coach. Data collected also represented four cohorts and spanned the years between 1997-1998 and 2005-2006 (Lockwood et al., 2010).

With a 95% confidence level, Lockwood et al. (2010) reported that on average across grades six through eighth, a state-funded coach was associated with “statistically significant improvements in average annual gains for 2003 and 2005 cohorts (p. 381). Grade eight results showed a positive and significant effect in three cohorts. The results of all four cohorts are listed below.

- 2003 cohort: effect was positive across all three grades and significant in grades seventh and eighth.
- 2004 cohort: effect was zero on average and across all three grades.
- 2005 cohort: all three grades showed positive and significant effects.

- 2006 cohort: the average effect is not significant; grade eight is the only grade that showed a significant and positive effect.

As reported, Lockwood et al.'s (2010) data from the 2006 cohort was the most creditable due to the use of data included in 1998 which provide pre-treatment data points and allowed for the identification of trends.

The model used in Lockwood et al.'s (2010) study expressed growth over time, so one can use the 2003 cohort, for example, to predict growth in 2006. According to the data, the 2003 cohort experienced an effect size of coaching of .06 on reading achievement. Using this model, one could estimate that achievement would be .24 standardized units higher than in schools that did not have a coach. While these results are promising, Lockwood et al. (2010) cautioned that the lack of consistency in evidence across cohorts and across grade levels suggested that coaching may be a "popular intervention," but it is not a "panacea for all schools" (p. 383) and that coaching may be more impactful in low performing schools with coaching implemented over a number of years.

A second study that examined coaching and student gains at the elementary level investigated the work of Reading First and its coaches. The purpose was to understand how coaches distributed their time and the rationale they gave for how their time was spent (Bean et al., 2010). For the purposes of this literature review, the focus was on the relationship between what coaches do and student achievement (Bean et al., 2010). As one of three research questions, Bean et al. (2010) examined what relationships exist among coaching qualifications, coach activities, and student achievement.

To analyze the variable of coaching activities, the schools were divided into two groups based on the amount of individual and group coaching sessions, using a mean split. This type of

comparison allowed the researchers to look at student achievement between schools with more and less coaching time with teachers. The DIBELS Oral Reading Fluency scores were compared in the fall for each group to establish a baseline. There was no significant difference in scores between the two groups in the fall. There was, however, a significant difference between the two group's scores' in end-of-year achievement in first and second grade. Schools in which the coaches spent more time coaching had a significantly greater percentage of students scoring as proficient in the Terra Nova in first and second grade and a significantly smaller percentage of students scoring at risk in the same grades (Bean et al., 2010).

In describing the limitations of interpretations in this study, the researchers conceded that although three weeks of coaching log data, coaching diary data proved insightful, they only provided a brief snapshot of how coaches spent their time. Second, the relationships between more or less coaching time and student gain scores were not causal. Bean et al. suggested the use of future experimental design to explore any causal relationship between the complex variables of “what coaches do, how teachers respond, and what students learn” (p. 111).

Specific to research in elementary schools, Elish-Piper and L'Allier (2010, 2011) provided a comprehensive study broken into two sections, each evaluating components of coaching and student achievement in grades Kindergarten through third grade. Their combined research provided mixed results as to the relationship in terms of gain scores per grade most impacted and total variance accounted for between coaching categories and gain scores.

Elish-Piper and L'Allier (2010) focused on a school district that received federal dollars in the form of a Reading First grant, the predecessor of Early Reading First. Their study reviewed similar lines of inquiry and asked two questions of a group of Kindergarten and first grade classrooms; 1) In what areas do literacy coaches primarily engage? and 2) What is the

relationship between literacy coaching and student reading achievement? Using descriptive statistics, hierarchical linear modeling, and multiple regression modeling, Elish-Piper and L’Allier were able to speak to the differences in the specific types of coaching (i.e., demonstration teaching) versus the content of interaction (i.e., comprehension, shared reading) and their relationship to student gains.

The results from this particular model of analysis provided important and useful information not only for the specific Reading First grant but also for the practice of literacy coaching for Kindergarten and first grade classrooms (Elish-Piper & L’Allier, 2010). A literacy coaches’ time was divided into two categories; time spent working directly with teachers (accounting for about 53%) and time spent in other duties such as planning meeting for PD workshops and organizing materials (accounting for about 47%). The researchers found that of the three top types of coaching from the coaches’ logs, modeling, observing, and conferencing accounted for 22.01% of the five literacy coaches’ time. The results showed that time spent working directly with teachers had the greatest gains in student reading achievement. Although findings from this study identified four specific coaching activities that predicted student reading gains, these aspects of coaching did not account for all the variation in student gains at the teacher level; therefore, future research is needed to identify other coaching activities and content areas that will predict student gains (Elish-Piper & L’Allier, 2010).

Elish-Piper and L’Allier (2011) built on their earlier research and examined the relationship between the amount, time, and content of literacy coaching by looking at Kindergarten through third grade reading gains. The two most relevant research questions were

- 1) Does the amount of time literacy coaches spend working directly with teachers predict student

reading gains in the classroom where they coach? and 2) Do specific literacy coaching activities and the specific content of literacy coaching predicts student reading gains?

In response to the first question involving coaching time spent working directly with teachers and student gains, the results showed some evidence to suggest that time, as measured by the literacy coaching log, may be related to student reading gains as measured by DIBELS. Total coaching hours were reported as a significant predictor at the second grade level and approached significance at the Kindergarten level (Elish-Piper & L'Allier, 2010).

The results addressing the second question of the most predictive coaching strategies of student reading gains showed four activities, conferencing, administrating and discussing assessments, modeling, and observation. These four activities were significant predictors of gains in one or more grades. It is noteworthy that none of the coaching variables proved to be predicative of student gains in the third grade. Also interesting to note is that no combination of variables accounted for more variance than any single-variable model (Elish-Piper & L'Allier, 2010).

The Institute of Education Sciences (IES) and the National Center for Education Evaluation and Regional Assistance (NCEE) also weighed in on the research question of professional development but found little effects of coaching to student achievement. The IES and NCEE commissioned the study, *The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement* (Garet et al., 2008). The DOE and IES study was designed to provide an analysis of the impacts of two different types of professional development, a teacher institute series (Treatment A) or a teacher institute series plus coaching (Treatment B), on the intermediate outcomes of teacher knowledge and teacher practices, and the ultimate outcome of student achievement. The study design did not permit an analysis of the

causal links of this theory of action, but did allow for measures of association with student achievement.

The study included 90 schools in 6 districts with 270 second grade teachers divided into 2 treatment groups and 1 control group.

- Treatment A consisted of a content-focused teacher institute series that began in the summer and continued through much of the school year (48 hours of professional development).
- Treatment B included the same institute series plus in-school coaching (on average 60 hours per teacher).

The 90 schools were randomly assigned to the 3 groups. A variety of data were collected from 2005-2006 as the intervention year and 2006-2007 as the follow up year, and several outcome measures were constructed. The first outcome measure included teachers' knowledge about reading instruction. The second measure examined teachers' use of research-based instructional practices. Students' reading achievement accounted for the third outcome measure.

The impact of these two types of professional development on teacher knowledge, classroom instruction, and student achievement produced three results (Garet et al., 2008).

- Teacher knowledge: There were positive impacts on teacher's knowledge of scientifically based reading instruction and on one of the three instructional practices in the classroom
- Classroom Instruction: The added effect of the coaching intervention (Treatment B) on teacher practices in the implementation year was not statistically significant.

- Student Scores: Neither Treatment A (teacher institute series) nor B (series plus coaching) interventions resulted in significantly higher student test scores at the end of the one-year treatment
- There were no statistically significant impacts on measured teacher or student outcomes in the year following the treatment.

Three possible hypotheses were identified that could help explain why changes in teacher knowledge and practice in the implementation year did not translate to impacts in student achievement. They included considerations of mobility, appropriate predictors of student achievement, and the size of the change in knowledge and practice needed to produce a change in student achievement.

Review of Studies Specific to the Field of Early Childhood Education

The research in the early childhood setting, similar to that of elementary grades studies, showed mixed results in the impact of coaching on child outcomes. The variations in studies and implications prove the need for further research in field of early childhood education. Four major studies were included and outlined for this review of literature, one focusing on a Head Start program, two from a multiple part study focused on an Early Reading First program, and the Final Evaluation of Early Reading First: Final Report to Congress.

Shidler (2009) explored the relationship between coaching and student gains in a Head Start classroom, serving four year old children. The study explored coaching, teacher efficacy in content instruction, and child outcomes and/or achievement. Teacher efficacy was based on “the teacher’s ability to see him/herself as capable of providing instruction within a content area and for the instruction provided to impact on student achievement” (Shidler, 2009, p. 453). The research question asked if more time coaching spent with teachers in the classroom resulted in

higher child outcomes. Shidler (2009) hypothesized that there will be a correlation between numbers of hours spent coaching (for the goal of teacher efficacy) and child outcomes, as measured by Alphabet Letter Identification and the Peabody Picture Vocabulary Test-III (PPVT-III).

In Year 1, teachers attended a 40 hour college course in emergent literacy. After each session, the coach visited the classroom to reinforce the coursework as well as model the practices. In Year 2, the coaches worked with teachers in their classrooms between six and ten hours a week for a nine month frame. In Year 3, coaches spent more time in the classroom than in Year 1, but less than in Year 2 and included various curriculum areas.

At the end of Year 1, Shidler (2009) found a significant correlation between the results of letter recognition skills and hours spent coaching the teachers. The classrooms that received higher coaching hours were more likely to see increased child outcomes in the same category. Years 2 and 3 were ranked through a system of ranking, including 1) the hours spent coaching, 2) child gain scores on the PPVT-III and 3) letter identification scores. Each classroom was given a score from first through ninth in each of the three categories. There was no significant correlation between rank per classroom and letter identification. There was also no significant correlation between gain scores and hours coaching per classroom to their ranking.

Teacher education was a suspected moderating variable. To address this additional question, a regression analysis was performed using hierarchical linear modeling. This analysis led to the conclusions that neither teacher education nor years of teaching experience were statistically significant to either letter recognition or gains scores. Overall, Shidler (2009) concluded that "a more focused, honed approach to coaching teachers in enhancing child outcomes in specific measures was more effective" (p. 459).

As in other educational arenas, research exists examining the impact of professional development in teacher knowledge and improving the quality of early language and literacy practices in early childhood settings. Neuman and Cunningham (2009) examined the impact of professional development in the form of coursework and coaching and their impact on teacher knowledge and practice. Their methodology included the use of three groups; Group 1 receiving coursework only, Group 2 receiving coursework and in class coaching and, Group 3 acting as a control.

Neuman and Cunningham (2009) were able to isolate the impact of the variables of coursework and coaching. The coursework included a 45-hour 3-credit hour course in language and literacy taught from professors from a local community college. The coaching model was diagnostic and prescriptive with a focus on supporting teachers in applying strategies to improve children's outcomes in language and literacy. The elements of the model included on-site coaching, support that was balanced and sustained, promoted teacher reflection, interaction that was built on a trusting relationship and was highly interactive, included corrective feedback, and focused on priorities and the development of action plans for children. The combination of these elements followed this cycle: coaches prompted teacher reflection and goal setting, supported teachers as they identified outcomes and the strategies necessary to reach them, and collaboratively developed an action plan for implementation of the new practices. These lessons then became the source for further reflection the following week.

Neuman and Cunningham (2009) hypothesized that professional development that included a combination of coursework and coaching would have a greater impact on improving teacher knowledge, practices, and strategies. Their results indicated no statistical differences in teacher knowledge between the three groups. There were, however; statistical significant

improvements in teacher practices between teachers who received both coursework and coaching (Group 2) and those receiving coursework only (Group 1). These findings indicate that professional development in the form of course work alone had “negligible effects on improvements in quality practices” (Neuman & Cunningham, 2009, p. 532).

The significance of this type of research lies in the light it sheds on how different types of professional development, including early childhood language and literacy practices, affect teacher knowledge, practices, and strategies; therefore, highlighting coaching as an important tool. Neuman and Cunningham (2009) reported their surprise at their findings that stand in contrast to Dickinson and Caswell (2007) who reported professional development alone resulted in significant gains in teacher practices. Neuman and Cunningham (2009) noted several possible factors contributing to the failure to replicate the results from Dickinson and Caswell, notably the fact that in Neuman and Cunningham’s (2009) study, the teachers attended the professional development sessions alone, without their administrators. Other noteworthy considerations included the quality of the coursework and the linkages between theory and practice.

In 2010, Neuman and Wright built on earlier research to address gaps they had previously identified. Specifically, their research would not determine if coaching was the change agent or if it was the combination of coaching plus coursework that provided for significant improvements in teacher practices. In the 2009 study, coursework was 45 hours while coaching included 64 hours; the difference in results could have been a dosage issue. The 2010 study asked four questions to examine these gaps including; are their differences between the forms of professional development on improvement of teacher knowledge and how might a smaller dosage influence professional development? In this study, Group 1 consisted of coursework only, Group 2 involved coaching only, and Group 3 acted as the control.

In regard to teacher knowledge, results from this study were comparable to Neuman and Cunningham's research in 2009. Neither treatment group, coursework, nor coaching significantly outperformed the control group in terms of teacher knowledge. In addition, it appeared from analysis that neither group treatment outperformed the other, nor the coursework or the coaching improved teacher knowledge. Using the Early Language and Literacy Checklist Observation to measure improvements in teacher practices, the results indicated statistically significant differences between group 2 and groups 1 and 3, illustrating that coaching alone impacted teacher practices. There was also no statistical difference between group 1 (professional development) and group 3 (control). In addition, follow up analysis of covariance tests pointed out these improvements in teacher practices were maintained five months later.

Jackson et al. (2007), in the *National Evaluation of Early Reading First: Final Report to Congress*, assessed the impact of ERF on teachers and classroom practices, by examining outcomes for teacher knowledge and skills and the general quality of the preschool environment. The quality of language, early literacy, and child-assessment practices and environments were also assessed. For teacher knowledge and skills, the researchers found that overall ERF had "positive impacts on the hours of teachers' professional development" during the 12 months preceding the survey and that it "increased the proportion of teachers receiving professional development through mentoring" (Jackson et al., 2007, p. 22). The study examined the measures of teacher instructional practice and the classroom environment specific to language and emerging literacy. According to the study's post survey, ERF had "*pervasive* impacts on the general quality of the preschool classroom" (p. 23, emphasis in the original). Most interestingly, ERF demonstrated impact on oral language use by both the lead and assistant teachers, teachers' book-reading practices, and teachers' use of expressive voice, phonological awareness activities,

and print and letter knowledge materials. However, researchers found overall that ERF had a “statistically significant positive effect on children’s print and letter knowledge but no statistically discernible impact on phonological awareness or oral language” (Jackson et al., 2007, p. 24).

These studies and their findings extending from the preschool through the elementary years prompt many interesting questions and platforms for future consideration and investigation. Researchers, practitioners, and policymakers focused on the field of language and literacy should consider the many variables involved in the complex nature of professional development and its’ relationships to child outcomes, teacher practices, and changes in the classroom environment (Garet et al., 2008).

Chapter 3 presents the methodology used in this study to answer the research questions, 1) What are the relationships between coaching hours and children’s literacy outcomes? and 2) What is the relationship between coaching hours and teacher practices and/or changes in the preschool classroom environment?

CHAPTER 3

METHODOLOGY

Research Questions

The literature indicates that individualized and sustained work with teachers that provides guidance and feedback on implementation of evidence-based practices in their own classrooms is superior to one-time workshops in improving teacher quality (Powell et al., 2010), yet the question remains: To what extent are coaching and child outcomes, teacher practices, and/or changes in the classroom environment related, given the assumptions previously described?

Intensive and sustained forms of professional development such as coaching with teachers are of particular interest; however, this evidence-based knowledge has been “widely ignored in early childhood programs” (Powell et al., 2010, p. 299). To add to the research and evidenced-base of this practical coaching conversation, the following questions were addressed in this study.

- 1) What are the relationships between coaching hours and children’s literacy outcomes?
- 2) What is the relationship between coaching hours and teacher practices and/or changes in the preschool classroom environment?

Research Design

According to the application “Local Non-Profit (LNP) Narrative_Education,” the document submitted for consideration and selection for the FY 2009 ERF grant cycle, the methods to evaluate achievement and validity/reliability data included a mixed-methods approach that “incorporates qualitative and quantitative data to track progress, provides ongoing

feedback for quality improvement and program refinement, and determines the effects of the program on targeted children and teachers” (Local Non-Profit, 2009, p. 36). GRANT was designed as a quasi-experimental matched sample pre to posttest design, repeated annually, and used to compare outcomes between participating children in an intervention group and a comparable group non-participating children. LNP (2009) believed that literacy coaching was an “essential piece to improving the instructional practices of early childhood teachers, and in turn, children’s language and literacy skills (p. 36). The evaluation design was expected to examine whether coaching was a critical component to the professional development of early childhood professionals and if child outcomes improved as a result of this practice (LNP, 2009).

Early Reading First

This chapter presents the methodology employed to answer the research questions. Due to the specific data set used in this study from a federally funded Early Reading First grant, an overview of Early Reading First and its location within a large urban area is necessary to understand more fully the study’s context and implications. The overview is presented here.

Early Reading First (ERF) was an initiative born from the No Child Left Behind (NCLB) Act of 2002. The overall purpose of the Early Reading First initiative was to “prepare young children to enter Kindergarten with the necessary language, cognitive, and early reading skills to prevent reading difficulties and ensure school success” (U.S. Department of Education, 2012, ¶ 3). The initiative was born in response to the growing concern that children entering Kindergarten did not have the necessary foundation in language and literacy exposure to benefit more fully from a formal Kindergarten instruction and a school experience.

The Government Performance and Results Act (GPRA) oversee the performance of Early Reading First programs and its grantees. The GPRA requires each agency and program “to set

targets for its performance, measure progress toward those targets, and report on whether or not the targets have been met, and describe future strategies for striving toward those targets” (U.S. Department of Education, 2009a ¶ 1). According to the Department of Education: Archived Performance website ¶ 3, the Secretary of Education established the following four measures for evaluating the overall effectiveness of the Early Reading First program for FY 2009:

- The cost per preschool-aged child participating in Early Reading First programs who achieves a significant gain in oral language skills after each year of implementation
- The percentage of preschool-aged children participating in Early Reading First programs who demonstrate age-appropriate oral language skills after each year of implementation
- The average number of letters Early Reading First preschool-aged children are able to identify after each year of implementation
- The percentage of preschool-aged children participating in Early Reading First programs who achieve significant gains in oral language skills after each year of implementation
- The Early Reading First teachers' average score on the Literacy Environment Checklist on the Early Language and Literacy Classroom Observation (ELLCO) Toolkit after each year of implementation (FY 2008 only)

The process of measuring these goals was designed to “improve program management,” and to help Congress, the Department of Education, the Office of Management and Budget, review a program's progress toward its goals (U.S. Department of Education, 2009a ¶ 1). Significant gains, in this situation, were defined as four or more standard score points between pre and posttesting. To date, only information from 2004-2007 was reported on the ed.gov website.

This dissertation recognized and acknowledged that one item of ERF’s effectiveness measure focused on cost of significant student gain, one item specified a specific skill

acquisition, and one item evaluated teacher performance. Of the fourth and fifth effectiveness measures, one item measured the percentage of students who demonstrated progress toward a predetermined bench mark, and one item specifically targeted significant student gains, in this case in oral language. The research and analysis for this study elaborated, in more depth, on two of these items, specifically:

- The item studying the percentage of students who achieved significant gains in oral language by teasing out the types of coaching activities and number of coaching hours from the coaching logs in relation to the gains to determine what relationships exist between coaching hours and student achievement
- The ELLCO scores from teachers that received multiple years of ERF professional development to understand the extent to which coaching hours and changes in teacher practices and/or environments are related over time

Context Specific Early Reading First

According to the document, Archived FY 2009 Grantee Abstract from the Department of Education, available through the DOE website, GRANT design “was based on our [LNP] previous success as an ERF grantee and will incorporate the use of a literacy-focused, research-based curriculum, Opening the World of Learning (OWL) and a curriculum supplement, BookFlix” (p. 10). LNP program goals included developing key early literacy skills, integrating high-quality professional development, use valid and reliable assessments, designing a high-quality environment, and providing a smooth school transition (U.S Department of Education, 2009b, ¶ 3).

Coaching Model:

The National Evaluation of Early Reading First in 2007 suggested that professional development be “continuous, intensive, and classroom focused” (Jackson et al., 2007, p. 27). Although a specific coaching model is not given, professional development that included coaching as a strategy was encouraged. The U.S. Department of Education (2003) also listed coaching as an example of professional development based on scientifically-based reading research. Examples of coaching, as defined in the document outlining guidance to Early Reading First, included “demonstration by the coach of effective strategies, and coaches’ observation of teachers’ instruction followed by discussion and reflection on the effectiveness of instructional strategies and how they support student progress” (U.S. Department of Education, 2003, ¶ 3).

The common sense approach described based on experience seemed to be the case with GRANT’s coaching model. In its more specific and localized setting, GRANT provided literacy coaches to support classroom teachers as one professional development strategy. These coaches were assigned to specific classrooms and worked in classrooms approximately four days a week, with one day reserved for planning and administration work. Responsibilities for the coaches during their classroom days included “observation, coaching, and providing feedback to teacher and assistants as they integrate curricula and materials and apply the knowledge base to classroom practice (St. Cyr, 2012, p. 17). Time for coaches focuses on sustaining learning and promoting reflection.

Figure 3 illustrates the LNP coaching model. The content of Figure 3 provides guiding questions for coaches and illustrates the relationship between the pre conference, observation, and post conference (St. Cyr, 2012).

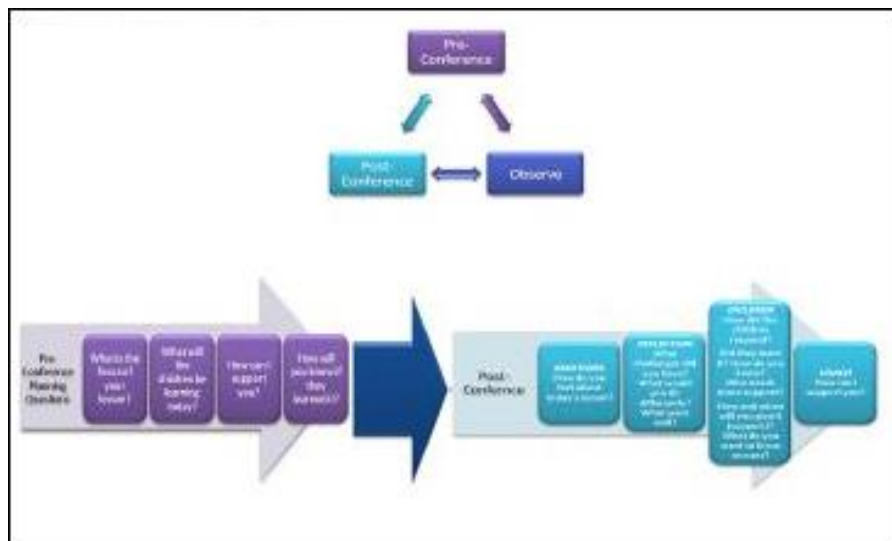


Figure 3. Local Non-Profit Coaching Model. (St. Cyr, 2012).

In addition to this coaching model (see Figure 3), the original grant document from LNP explained that in the area of high quality instructional strategies, “professional development will focus on providing integrated, sequenced instruction using advanced strategies in intentional, explicit, and systematic instruction as well as scaffolding and sufficient practice” (LNP, 2009, p. 4).

Specific coaching practices:

The coaching subcategories were pulled from a coaching document that GRANT’s Coaching Director created, refer to Figure 1. At the beginning of the third year of implementation, the coaches discussed the subcategories and made their final suggestions. All four coaches agreed to code their time logs using the description outlined in the document.

Participants

The intervention group consisted of 206 children in 15 classrooms and 15 lead teachers located in 6 different sites. This group received the training and materials for the Opening the World to Learning (OWL) curriculum, access to a literacy coach on a regular basis, and the

opportunity to attend (four) six-hour professional development sessions offered by an external provider, (three) six-hour professional development sessions offered by the coaching staff, a six day Challenging Teacher Institute, and a one day conference.

The comparison group consisted of 85 children in 6 classrooms and 6 lead teachers located in the same site. This group received the training and materials for the Opening the World to Learning (OWL) curriculum, and the opportunity to attend (four) six-hour professional development sessions offered by an external provider, (three) six-hour professional development sessions offered by the coaching staff, a six day Challenging Teacher Institute, and a one day conference. The comparison group did not receive access to a literacy coach.

Student Information:

At the beginning of each school year, an introductory letter explaining the purposes of the grant was distributed to each student. Each child's parent/guardian had the option of accepting or declining their child's participation in the grant, but declining additional evaluation did not affect the children's participation or status in their classroom or site. Included in this letter was a request for demographic information, such as gender and ethnicity. At the GRANT Director's request, this information was listed as "optional;" therefore, not every child enrolled in the classroom was assessed and not every child assessed provided complete demographic information.

Both the intervention and comparison classrooms had comparable proportions of males and females, with both sites containing a slightly higher percentage of females. The intervention classes contained 126 males (39%) and 196 females (61%), while the comparison classrooms contained 55 males (44%) and 70 females (56%). Additionally, in the intervention group only six children (2%) self-reported having an Individual Education Plan (IEP), and no children self-

identified being English Language Learners (ELL). In the comparison group, no children self-identified as having an IEP or as being an ELL student.

There was no available ethnicity data on children in either the intervention or the comparison group, although “we believe there are no real differences in ethnicity between the groups based on prior years” (St. Cyr, 2012, p.14). As a footnote in the final evaluation, the evaluator noted that there were numerous discussions with project staff and data collection associates, but the reasons for the missing data remain unclear. In years 1 and 2, this information was collected. It is important to note that not all children in the intervention or comparison classrooms were assessed; and therefore, they were not included in the data set. Reasons for the discrepancy between total classroom attendance and total children in the data set range from a family’s decision not to participate in the assessment component of GRANT to children absenteeism during screening periods. These real life teaching situations also account for the differences in the number of children screened per assessment.

Teacher Information:

Table 1 provides information on teacher’s ethnicity, age, highest level of education, and number of years teaching. The intervention group equaled an *n* of 44; the comparison group equaled an *n* of 12.

Table 1

Demographic Information for All Teachers

Demographic		Intervention (n=44)	Comparison (n=12)
Ethnicity	African American	24 (89%)	12 (100%)
	Caucasian	2 (7%)	0
	Hispanic	0	0
	Other	1 (4%)	0
Age	20-30	11 (36%)	1 (11%)
	30-40	11 (36%)	4 (44%)
	40-50	6 (19%)	2 (22%)
	Over 50	3 (10%)	2 (22%)
Highest Education	High School/GED	7 (23%)	3 (25%)
	AA/AAS	10 (33%)	5 (42%)
	BA/BS	12 (40%)	4 (33%)
	Masters	1 (3%)	0
Years Teaching	1-5 Years	10 (36%)	0
	6-11 Years	9 (32%)	3 (38%)
	12-19 Years	4 (14%)	4 (50%)
	20 Years Plus	5 (18%)	1 (13%)

Note: Information includes lead teachers, assistant teachers, and paraprofessionals in the classroom. The table was recreated from Table 4 in the Year 3 Annual Evaluation Report (St. Cyr, 2012).

Coach Information:

Information on coaching demographics included ethnicity, highest level of education achieved, years teaching in a classroom, and years taught in program management and/or leadership positions. These data points were collected at the beginning of year 3 by LNP as part of another project. LNP gave permission to the researcher for the information to be included for this dissertation's research. All four coaches self-reported as African American and as having a Master's degree. As part of the same survey, 25% of the coaching staff have taught in the classroom between 1-5 years and 75% have taught between 12-19 years. When asked about years of experience in program management and/or leadership positions, 50% of the coaches have between 0-5 years experience in management and/or leadership and 50% have between 6-11 years experience. Other experiences include adjunct professor (ECE), trainer, and literacy coach.

Site Information:

In FY 2011-2012, GRANT operated in seven different sites, six comprised the intervention group, and one represented the comparison group. These 7 sites included 10 Pre-K classrooms, 4 Preschool classrooms, and 1 mixed-age classroom, totaling 15 classrooms and 322 students. The 1 comparison site housed 6 Pre-K classrooms and a total of 125 students. Not all students enrolled in the classrooms granted GRANT permission to collect student data and of those who granted permission, only students with pre and posttest scores available were included in the analysis.

Table 2 describes the sites in more detail, including the number of classrooms, number of children pre-tested, the sites' funding source, the number of days a year the program operated for

the children in the GRANT classrooms, the sites' free/reduced lunch status, their previous curriculum, and the site's FY 2011-2012 child attrition rate for the GRANT classrooms.

Table 2

FY 2011-2012 GRANT Classrooms

	# of classrooms	# of children pretested	Funding Source	Free/Reduced Lunch Rate	Previous Curriculum	Child Attrition Rate
Intervention						
Bolton Academy*	1 Pre-K	19	GA Lottery			0%
Cook Elementary*	1 Pre-K	15	GA Lottery		Blue Print for Literacy	20%
Dunbar Elementary*	1 Pre-K	18	GA Lottery	99%		11%
JF K Middle School*	1 Pre-K	16	GA Lottery			13%
Our House	1 Pre-K, 1 Preschool, 1 Mixed-Age	49	Grants and Scholarships No cost to families	100%	Creative Curriculum	31%
Renaissance Learning Center	5 Pre-K 1 Preschool	107	Pre-K, GA Lottery Preschool, Sliding Scale Tuition	85%	Creative Curriculum	12%
Dunbar-Sheltering Arms	2 Preschool	34	Sliding Scale Tuition	99%	Sheltering Arms Preschool Curriculum	6%

	# of classrooms	# of children pretested	Funding Source	Free/ Reduced Lunch Rate	Previous Curriculum	Child Attrition Rate
Comparison						
KiDazzle	6 Pre-K	98	GA Lottery	92%	Creative Curriculum	12%

Note: Four Pre-K classrooms fall under the authority and leadership of the Atlanta Public Schools Pre-K office; Bolton Academy, Cook Elementary, Dunbar Elementary, and JF Kennedy Middle School.

Measures

The children were evaluated with three assessments, the Peabody Picture Vocabulary Test, Fourth Edition (PPVT), the Expressive Vocabulary Test, Second Edition (EVT), and the Phonological Awareness and Literacy Screening (PALS Pre-K) (St. Cyr, 2012). The teachers' classroom practices and changes in the environment over time were measured with one assessment, the Early Language and Literacy Classroom Observation (ELLCO).

The Peabody Picture Vocabulary Test, Fourth Edition (PPVT):

The Peabody Picture Vocabulary Test measure is used to assess vocabulary acquisition by measuring understanding of the spoken word (receptive language) in standard American English. The assessment is used as a brief achievement measure of receptive language and as a general screening measure for verbal ability. The PPVT scale is a norm-referenced, wide-ranged instrument available in two forms that are administered individually. Each form consists of training or practice items and 228 test items. Each test item consists of four full-color pictures. The examiner says a word and the child responds by pointing to the picture that best illustrates

that word's meaning. The items broadly sample words from 20 content areas and parts of speech (nouns, verbs, attributes) (Dunn & Dunn, 2007).

The PPVT can be used to screen ages 2 years 6 months old through 90 years and older. Administration is untimed, but typically takes between 10 to 15 minutes. According to the technical manual, there are 15 primary purposes and uses. The PPVT measures responses to vocabulary instruction, contributes to assessment of preschool age children, helps in the detection of language, and screens for verbal development (Dunn & Dunn, 2007).

Scores/or subscales that emerge from scoring the PPVT include a nationally represented sample of a standard score of 100, with a standard deviation of 15. One standard deviation below the mean is 85; therefore, 85+ is considered acceptable or average, 80-84 is considered low average, and 20-79 is considered below average. The PPVT is also normed by age and grade. According to Dunn and Dunn (2007), other indicators collected from the national samples include sex, race, ethnicity, geographic region, socioeconomic status/parent education, community size, and special population information. Reliability and internal consistency were assessed with Guttman split-half reliability measures by age were .94 and .94 on Forms A (fall) and B (spring), and by grade were .95 for Form A and .94 for Form B. The test-retest reliability mean for age was .93.

According to Campbell, Bell, and Keith (2001), in an educational screening instrument, predictive validity is the most important type of psychometric validity. This particular type of validity refers to an assessments ability to identify those with the concern of interest, in this case oral language delays, while excluding those without this concern. In terms of validity, the PPVT was assessed and the stimulus words were chosen from a review of over 12 published reference works and represented 20 content areas. The construct validity was correlated with five different

instruments, including the PPVT-III and the EVT. The observed pattern of correlation was consistent with the expected pattern of a valid vocabulary measure. Criterion-related validity was correlated with the same five instruments. The correlation between the PPVT and the PPVT-III was .84 and the PPVT and EVT was .82.

The Expressive Vocabulary Test, Second Edition (EVT):

The EVT measure is used to measure expressive vocabulary and word retrieval (vocabulary acquisition) of the spoken word in standard American English. The EVT is also a norm-referenced instrument that assesses expressive vocabulary and word retrieval in this case, for children. Like the PPVT, there are two parallel forms that are administered individually. LNP uses Form A in the fall and B in the spring. Each form contains sample or practice items and 190 test items that are arranged in increasing difficulty. During administration, the examiner presents a picture and asks a scripted stimulus question. The child is expected to answer with one word that provides an acceptable level, answers a specific question, or produces a synonym for the word in the picture (Williams, 2007).

The EVT can be used to screen children ages 2 years 6 months through adults age 90 and older. This assessment is also untimed and takes an average between 10 to 20 minutes, depending on the age of the child and existing vocabulary knowledge. Within the ERF context, there are several primary uses or purposes that are most useful and contribute to the assessment of preschool children. The EVT measures word retrieval, monitors growth across a broad time span, and can be used for direct comparisons between expressive and receptive vocabulary skills with the PPVT scale (Williams, 2007).

Scores and/or subscales that emerge from scoring the EVT include a nationally represented sample of a standard score of 100, with a standard deviation of 15. One standard

deviation below the mean is 85; therefore, 85+ is considered acceptable or average, 80-84 is considered low average, and 20-79 is considered below average. The EVT is also normed by age and grade. Other indicators collected from the national samples include sex, race, ethnicity, geographic region, socioeconomic status/parent education, community size, and special population information. Reliability and internal consistency were assessed with Guttman split-half reliability measures by age were .94 on Form A (Fall) and .93 on Form B (Spring) and by grade were .93 for both Forms A and B. The test-retest reliability mean for age was .95.

In terms of validity, the EVT was assessed for content validity; the selection of words was chosen from a review of over 9 published reference works and represented 20 content areas. The construct validity was correlated with five different instruments, including the PPVT and EVT. The observed pattern of correlation was consistent with the expected pattern of a valid vocabulary measure. Criterion-related validity was correlated with the same five instruments. The correlation between the EVT and PPVT was .82 and the EVT-I and EVT was .81.

Despite the PPVT's wide-spread use as a screening instrument for verbal ability for programs receiving federal funding, Restrepo, Schwanenflugel, Blake, Neuharth-Pritchett, Cramer, and Ruston (2006) "strongly caution practitioners in the use of the PPVT-III for verbal ability estimates or screening and for identification of language disorders when assessing African American children whose mothers have less than a high school education (p. 25). In the current research sample, no specific data on this variable were collected; nevertheless, caution should still be taken seriously in situations in which this information is most relevant.

Restrepo et al. (2006) reported interesting findings in their research on the PPVT-III in response to questions involving test bias by ethnic group, parent education, and gender, as well as discrepancies by parent education and ethnic background. First, even though the PPVT-III

and the EVT were co-normed, the PPVT-III is a more difficult test, “particularly for African American children and, perhaps, for children whose mothers have only a high school education or less” (Restrepo et al., 2006, p. 23). Overall, these findings suggested possible bias against African American children and increasing levels of bias as a function of decreasing levels of maternal education levels (Restrepo et al., 2006).

Similarly, Campbell et al. (2001) also questioned the validity of the PPVT-III, in respect to accurately presenting the verbal ability of low income African American children. In their research, low socioeconomic status (SES) African American children scored more than one standard deviation below the mean of the PPVT-III standardization sample. Thus on average, the PPVT-III tends to underestimate both intellectual ability and scholastic achievement in this particular subset of the population (Campbell et al., 2001).

The Phonological Awareness and Literacy Screening (PALS Pre-K):

This screening tool is used to measure alphabet knowledge, phonological awareness, and print and word awareness in children. These skills are divided into six tasks, but GRANT only collected data for the following four items:

- Alphabet Knowledge (upper and lower case)
- Beginning Sound
- Print and Word Awareness
- Rhyme Awareness

The nature and purpose of the PALS Pre-K is to “guide teachers’ efforts at planning literacy instruction (Internizzi, Sullivan, Meier, & Swank, 2004). Only children who were four years old at the time of the pretest were assessed.

Scores and/or subscales that emerged from scoring the tool are included in Table 3.

These ranges were based on the spring developmental ranges as reported in the technical manual (Invernizzi et al., 2004).

Table 3

Scores and/or Subscales for the PALS Pre-K

	Acceptable	Borderline	Below Average
Alphabet Knowledge (Upper Case)	16+ (out of 26)	10-15	0-9
Beginning Sound Awareness	6+ (out of 10)	4-5	0-3
Print and Word Awareness	8+ (out of 10)	6-7	0-5
Rhyme Awareness	6+ (out of 10)	4-5	0-3

The reliability and internal consistency for the EVT was assessed using Guttman split-half reliability and Cronbach's alpha level. For both assessment measures, a score of 1.0 represents perfect reliability. In the pilot sample in 2003-2004, (six tasks of the PALS Pre-K, name writing, alphabet knowledge, beginning sound, rhyme, print and word awareness, and nursery rhyme awareness), all measures of internal consistency fell within acceptable ranges. Measured with Pearson correlation coefficients, Guttman's ranged between .94 and .75. Cronbach's alpha level ranged between .93 and .75. The internal consistency was also tested in the same pilot study. Five of six tasks were tested and all achieved correlations of .99 (Invernizzi et al., 2004).

Analyses of the pilot samples confirmed that PALS Pre-K measures a unitary trait, emergent literacy. This factor accounted for between 34% and 76% of the total variance in

children's scores and confirmed that construct validity had been established. Criterion-related validity refers to the relationship between the scores from the assessment and an external measure. Concurrent validity in the PALS Pre-K was measured against three independent assessments. All three showed a correlation, confirming criterion-related validity. Predictive validity was established through multiple regression analysis. Using 41 participants from the first pilot of the PALS Pre-K in the Spring of 2000, it was observed that the core task scores from a spring PALS Pre-K significantly predicted a large portion of the variance in the scores of a fall PALS Pre-K assessment ($R^2=.84$).

Invernizzi et al. (2004) offered two cautions about interpreting these correlations. First, the PALS Pre-K was designed as a diagnostic tool to help teachers to guide their literacy instruction while PALS K and PALS 1-3 were designed as a screening tool to identify children who would probably be in need of additional instruction. The second caution is one that should be considered in general when analyzing longitudinal data and involves the uneven nature of children's development (Invernizzi et al., 2004).

The Early Language and Literacy Classroom Observation (ELLCO):

This observation tool is designed to determine the effectiveness of classroom environments using a rated system of 1 through 5 over 5 areas (15 questions total), a score of 1 indicating deficient and a score of 5 indicating exemplary. The results can be used to strengthen the quality of a program and/or teaching practice and to improve young children's literacy outcomes.

The ELLCO measures five key literacy elements: classroom structure, curriculum, the language environment, books and book reading opportunities, and print and early writing supports (Smith, Brady, & Anastasopoulos, 2008). The ELLCO "explicitly values how materials

are used by teachers and children more than whether materials are merely present in the classroom” (Smith et al., 2008, p. 4). This specific tool gives educators and researchers a way to confidently examine features in a classroom specific to literacy by building anchor statements to each indicators, including descriptive indicators, and requiring evidence in each section as a foundation and rationale for a given score (Smith et al., 2008).

The ELLCO provides a common lens through which to examine and investigate literacy practices. Thorough training is required in order to conduct a consistent observation of the classroom that focuses on the evidence present; therefore, an evaluator should take careful consideration of the administration of this assessment (Smith et al., 2008).

The ELLCO should be used in Pre-K classrooms and the observer should allow at least 3.5 hours per classroom. The observer should plan to observe multiple activities, including large group, free-choice time, mealtime, and greetings/departures, and take notes in all 19 sections of the booklet, paying specific attention to the quality of the evidence. Data collectors should also participate in joint observations and calculate the interrater reliability level that is appropriate for the particular study. Follow up calibrations were also recommended (Smith et al., 2008).

The overall ELLCO max score is 95. The max scores and/or subscales that are present in scoring the ELLCO include: Classroom Structure (max of 20), Curriculum (max of 15), Language Environment (max of 20), Books and Book Reading (max of 25), and Print and Early Writing (max of 15). The quality rating in each item can range from one through five; deficient to exemplary. However, to obtain a score that allows for the ability to track progress over time, Smith et al. (2008) recommended dividing the sections into two subscales and use the scale of one through five with the number corresponding to the level descriptor (i.e., Level Two is

inadequate and Level Four is strong). After these sessions are totaled, they are divided by either 7 in subscale 1 or 12 in subscale 2.

- Subscale 1- [Classroom structure (1) + Curriculum (2)] / 7
- Subscale 2- [The Language Environment (3) + Books and Book Reading (4) + Print and Early Writing (5)] / 12

Reliability and internal consistency were assessed with Cronbach's alpha. The Classroom Observation total was .90, the General Classroom Environment was .83, and Language, Literacy, and Curriculum was .86. The interrater reliability between observers with training and appropriate supervision was 90%, scored with relative ease (within one point of each other on the rating scale).

The ELLCO was used in conjunction with the Classroom Profile. Convergent validity in the Classroom Environment score was found in the positive relationship between it and the Classroom Profile's Learning Environment subscale. Divergent validity was determined due to the absence of relationship between the Classroom Environment score and the Scheduling subscale of the Classroom Profile. Smith et al. (2008) believed that this total measures something qualitatively different and specific to the early language and literacy experiences and their classrooms. In terms of predicting child outcomes concerning classroom-related variance, the Classroom Observation accounted for 80% of between-class variance in vocabulary and 67% of between classroom variance in early literacy.

Data Collection

The PPVT, EVT and PALS Pre-K data were collected by an independent consulting agency and analyzed by a third-party evaluator as part of a longitudinal study of ERF. The ELLCO pre and post assessments were conducted by the same two coaches per classrooms. The

coaches did not assess their own classrooms. Data for Year 1 were collected in January 2010 (pre) and April 2010 (post). Data for Year 2 were collected September 2010 (pre) and April 2011 (post). Data for Year 3 were collected in September (2011) and April 2012 (post). Coaching logs for Year 3 recorded information from October (2011) through May (2012). Secondary analysis of archived data, children's measures, teacher practices, changes in the environment, and coaching logs, were used to address research question one. The same methodology was used to address research question two. Data sources included teacher practices and changes in the environment data from Year 1 (2009-2010), Year 2 (2010-2011), and Year 3 (2011-2012) and coaching logs from Year 3 (2011-2012).

Data Analysis

The analysis used to address these questions included an analysis of variance between coaching hours and 1) changes in child posttest scores and 2) changes in classroom ELLCO posttest scores. Correlation analyses were also run to determine the trend, the strength, and the direction of those relationships.

CHAPTER 4

FINDINGS

Introduction

This chapter details findings from a longitudinal evaluation of a local Early Reading First (ERF) program (GRANT) and specifically the data from Year 3 of that study. It is important to situate the findings from the current study in the context of the overall longitudinal study. Following an introductory summary of that larger study, findings from the Year 3 analysis addressing two research questions on coaching hours are presented. The research questions from the current study are a logical extension of GRANT'S analysis and focus more specifically on the relationship between coaching hours and child outcomes, teacher practices, and/or changes in the classroom environment.

Results from the Longitudinal Study of Early Reading First (GRANT)

Six guiding questions were examined as part of the GRANT study. Of the six original questions, three support the additional analyses for the current study: 1) What is the extent of teachers' participation in, and response to, professional development and coaching?, 2) To what extent do teachers increase use of SBRR (Scientifically Based Reading Research) strategies and curriculum?, and 3) How has professional development enhanced the knowledge, skills, and practices of GRANT teachers?

Variables examined in the overall GRANT project included teachers' participation in professional development sessions, attendance at the ERF Symposium, and participation in study groups from June 2011 to May 2012. Data from those analyses indicated each intervention

teachers averaged 45.9 hours of professional development training, while teachers in the comparison group averaged 22 hours. Professional development for teachers in the intervention group included support from a literacy coach, while no coaching support was provided for comparison teachers. Classrooms received an average of 45.3 hours of coaching over 9 months. Coaching support averaged 21.9 hours per teacher in the third program year (St. Cyr, 2012, pp. 30-31).

Data from the GRANT project suggested that all teachers in the intervention group rated their respective professional development experiences helpful, while a small percent (8-17%) of teachers in the comparison group disagreed. The final evaluation from Year 3 indicated intervention teachers reported their coaches were “knowledgeable, responsive, and beneficial to them and their classroom” (St. Cyr, 2012, p. 4). Measures were also taken on teachers’ comprehension and engagement in professional development. Teachers in the comparison group received comparable comprehension and engagement scores as the teachers in the intervention group.

Data from the GRANT study also assessed if teachers increased their use of Scientifically Based Reading Research (SBRR) strategies and the degree to which they implemented the prescribed curriculum, as measured by the OWL Fidelity checklist and the ELLCO tool. At the end of the third year of the study, 13% of teachers in the intervention group were implementing SBRR strategies at a high level and 87% at an intermediate level. No teachers in the comparison group were implementing SBRR strategies at a high level, performance ranged between the low to intermediate levels.

The GRANT study also assessed how professional development experiences enhanced the knowledge, skills, and practices of the GRANT teachers. On measures of a knowledge

survey of SBRR, the mean score for teachers in the intervention group decreased from pretest to posttest while the mean score from the comparison group teachers increased. At the end of Year 3, the evaluator reported that “compared to teachers in the comparison group, intervention teachers scored higher in measures assessing skills acquisition and classroom practices but not in a test of their knowledge of SBRR” (St. Cyr, 2012, p. 5). In the application of SBRR, as measured by the ELLCO Pre-K, intervention group teachers performed in the intermediate-high level and comparison group teachers performed in the low-intermediate level in year 3.

Overall, comparison teachers scored higher on measures of their knowledge of SBRR but their application of these strategies was between low to intermediate levels. Inverse findings were reported for intervention teachers as their application of SBRR was between intermediate to high; although, their knowledge of these skills decreased at the end of Year 3. Specific implications of these findings are presented in Chapter 5.

Finally, data from the GRANT evaluation indicated that all children, regardless of group assignment increased their skills in oral language, alphabet knowledge, phonological awareness, and print and word awareness. Children assigned to intervention classrooms slightly outperformed children from the comparison group in the areas of expressive vocabulary and rhyme awareness. While data from this GRANT evaluation provide direct measure of teacher and child performance, an additional secondary analysis delves more deeply into the unique contribution of coaching in understanding children’s outcomes. These results are presented in the next section of this chapter.

Results from Secondary Data Analysis

The purpose of this secondary analysis was to shed light on the relationship between coaching hours and child outcomes, teacher practices, and/or changes in the classroom

environment. To add to the evidenced-base of contemporary practical coaching conversations, the following questions were addressed in this study:

- 1) What are the relationships between coaching hours and children's literacy outcomes?
- 2) What is the relationship between coaching hours and teacher practices and/or changes in the preschool classroom environment?

Coaching was operationalized as an instructional practice and consisted of four distinct categories of behaviors or practices exhibited by coaches. Within Year 3 of the longitudinal study, 3851.3 hours of coaching were recorded which included both instructional and non-instructional coaching. The present study only focused on the instructional hours of coaching. As a result, a total of 1,122.95 instructional coaching hours were recorded in Year 3. Within this total, four subcategories of instructional practices were coded. These categories included Public Practice (85.05 hours, 7.57% of total), Pre and Post Conference (416 hours, 37.05% of total), Observation (403.40 hours, 35.92% of total), and Study Groups (218.50 hours, 19.46% of total). Of the total hours accounting for instructional and non-instructional coaching activities, literacy coaches spent 29% of their time engaged in instructional coaching activities.

Question 1: What are the relationships between coaching hours and children's literacy outcomes?

To address this first research question, a series of repeated measures analyses were conducted to examine the relationship of coaching hours on children's literacy outcomes, as measured by the PPVT, EVT, and the four sections PALS Pre-K. Specifically, these three literacy measures were examined from pretest to posttest to assess whether or not children whose teachers received coaching in their classrooms increased their scores over the intervention year on each measure. The first measure examined was the PPVT, a measure of receptive vocabulary.

On this measure, significant differences were found on the time analysis from pretest to posttest for students. No statistically significant differences were found in the interaction term or between the intervention and comparison teachers on this subscale. This time finding from pretest to posttest is consistent with the findings from GRANT analyses reported above.

The second measure examined was the EVT, a measure of expressive vocabulary. On this measure, significant differences were found on the time analysis from pretest to posttest for students. This, again, was consistent with the GRANT findings. No statistically significant differences were found in the interaction term; however, analyses indicated a significant difference between the intervention and comparison children on this subscale, suggesting that the effect of coaching support for the intervention children made a statistically significant difference in their relative performance on the EVT posttest measure.

The PAL Pre-K assessment was broken into its four individual sections for analysis; Upper Case letter recognition, Beginning Sounds, Print and Word Awareness, and Rhyme Awareness. For Upper Case letter recognition, there was a significant difference in the time analysis for children, but no statistically significant differences were found in the interaction term or between the intervention and comparison groups. For the Beginning Sounds and Print and Word Awareness sections, a similar pattern of findings was present. For both PALS Pre-K subtests, a statistically significant difference between children's pretest and posttest scores was evident, but no statistically significant differences were found in the interaction or group analyses. For the Rhyme Awareness section, two significant findings were evident. The time effect between pretest and posttest and the interaction term between Rhyme Awareness and condition were both significant; however, there was not a statistically significant difference for the group effect. This finding suggests that the time by group effect indicated a statistically

significant interaction meaning the children in the intervention group witnessed a steeper increase; although, the posttest scores between comparison and intervention groups did not suggest a final statistically significant difference in posttest scores. All means and standard deviations along with the specific values for the analysis of variance tests are found in Table 4.

Table 4

Results of Repeated Measures Analysis of Variance for the PPVT, EVT, and PALS Pre-K Measures

	Time 1 (pretest)		Time 2 (posttest)		F_T	F_{TxG}	F_G
	M	SD	M	SD	$\frac{p}{\eta^2}$	$\frac{p}{\eta^2}$	$\frac{p}{\eta^2}$
PPVT					29.56*	.52	.77
Intervention	91.97	15.39	95.27	13.93	.00	.47	.38
Comparison	89.81	15.57	94.12	12.28	.09	.00	.00
EVT					12.58*	2.28	3.93*
Intervention	94.77	15.05	97.64	13.10	.00	.13	.05
Comparison	92.10	12.11	93.26	11.34	.04	.01	.01
PAL Pre-K Upper Case					204.44*	1.03	.10
Intervention	12.94	10.05	20.78	7.32	.00	.31	.10
Comparison	13.10	9.95	19.90	7.75	.455	.00	.75
PAL Pre-K Beginning Sounds					148.47*	.11	.11
Intervention	4.01	3.83	6.93	3.54	.00	.74	.74
Comparison	4.07	4.02	7.16	2.95	.38	.00	.00
PAL Pre-K Print, Word Awareness					48.30*	2.49	1.40
Intervention	4.77	2.70	6.38	2.18	.00	.12	.24
Comparison	5.41	2.32	6.43	2.27	.17	.01	.01

	Time 1 (pretest)		Time 2 (posttest)		F_T	F_{TxG}	F_G
	M	SD	M	SD	$\frac{p}{\eta^2}$	$\frac{p}{\eta^2}$	$\frac{p}{\eta^2}$
PAL Pre-K					65.05*	12.15*	2.77
Rhyme Awareness	4.52	2.67	6.82	2.92	.00	.00	.10
Intervention	4.66	2.29	5.57	2.64	.21	.05	.01
Comparison							

Note: The maximum score for the PALS Pre-K sections are Upper Case letter recognition (26), Beginning Sounds (10), Print and Word Awareness (10), and Rhyme Awareness (10); values included for η^2 (eta squared) are .01 for a small effect, .06 for a medium effect, and .15 and higher for a large effect. *Statistical test is significant at the .05 level (2-tailed).

The PPVT, EVT, and PAL Pre-K scores were also examined as a function of hours of classroom support that a classroom received from coaching. To examine this question, a series of correlation analyses were conducted to examine strength and direction of the association. Correlations examined children's posttest scores for the PPVT, EVT, and PALS Pre-K and their relationship with total coaching hours per classroom per subcategory for the 15 intervention classrooms only. For the PPVT measure, two of the four correlations were both negative and statistically significant, indicating that as coaching hours in Public Practice and Pre and Post Conferencing increased, student scores on the PPVT decreased. However, coaching hours spent in Observations and literacy outcomes were significantly and positively correlated indicating as Observation increased so did children's PPVT scores. No statistically significant relationship was found between Study Groups and children's receptive vocabulary. Correlations between the PPVT posttest scores and coaching hours were as follows: Public Practice $-.18$ ($p = .01$), Pre and

Post Conferencing $-.18$ ($p = .01$), Coach Observation $.16$ ($p = .02$), and Study Groups $.04$ ($p = .53$).

For the EVT measure, two of the four correlations were significant, one positive and one negative. The correlational analysis suggested two findings; 1) as coaching hours for Pre and Post Conferencing increased, the EVT posttest scores decreased and 2) as coaching hours for Observation increased, so did the EVT posttest scores. No statistically significant relationship was found for Public Practice or Study Groups and the EVT. The correlations between the EVT posttest scores and coaching hours were as follows: Public Practice $-.05$ ($p = .46$), Pre and Post Conferencing $-.18$ ($p = .01$), Coach Observation $.15$ ($p = .02$), and Study Groups $.08$ ($p = .48$).

For the PALS Pre-K measure and its four sections, 16 correlations were conducted. Of those 16 tests, 75% of the correlations were negative (12 out of 16). However, only two were statistically significant and negative, suggesting that as coaching hours in Pre and Post Conferencing and Study Group increased, posttest scores in the Print and Word Awareness section decreased. There were no statistically significant positive correlations found. The correlations between PALS Pre-K posttest scores, as well as the PPVT and EVT, and coaching hours per subcategory are found in Table 5.

Table 5

Correlation Analysis for Coaching Practices and for the PPVT, EVT, and PAL Pre-K Measures

	Coaching: Public Practice	Coaching: Pre and Post Conferencing	Coaching: Coach Observation	Coaching: Study Groups
	<i>r</i> <i>p</i>	<i>r</i> <i>p</i>	<i>r</i> <i>p</i>	<i>r</i> <i>p</i>
PPVT	-.18*	-.18*	.16*	.04
Intervention	.01	.01	.02	.53
EVT	-.05	-.18*	.15*	.08
Intervention	.46	.01	.02	.48
PALS Pre-K				
Intervention	.01	.04	-.0	-.07
Upper Case	.95	.56	.29	.25
PALS Pre-K				
Intervention	.06	-.11	.09	-.12
Beginning Sounds	.40	.13	.25	.13
PALS Pre-K				
Intervention	-.04	-.18*	-.01	-.25*
Print and Word Awareness	.59	.02	.91	.00
PALS Pre-K				
Intervention	-.13	-.01	-.03	-.05
Rhyme Awareness	.08	.89	.66	.53

*Statistical test is significant at the .05 level (2-tailed).

Question 2: What is the relationship between coaching hours and teacher practices and/or changes in the preschool classroom environment?

To address the second research question, a series of repeated measures analyses were conducted to examine the relationship of coaching hours on teachers' practice at the classroom level. Specifically, five subscales from the ELLCO measure were examined from pretest to posttest to assess whether or not teachers who received coaching in their classrooms increased their ratings over time on the ELLCO measure. On the Classroom Structure subscale, no significant differences were found on the time analysis from pretest to posttest for classrooms. In addition, no statistically significant differences were found in the interaction term or between the intervention and comparison classrooms on this subscale.

The ELLCO measure was divided into a five-point scale per question, ranging from deficient with minimal evidence (1), inadequate with limited evidence (2), basic with some evidence (3), strong with sufficient evidence (4), and exemplary with compelling evidence (5). A description of classroom performance was determined by dividing the total score per section by five. Although no statistically significant differences on this subscale were found, it appears that classrooms had a moderately strong rating, an average of a score of four points per question with sufficient evidence, on this subscale of the ELLCO regarding classroom structure. Means and standard deviations along with the specific values for the analysis of variance tests are found in Table 6.

The same pattern of findings was observed on the Curriculum subscale where no statistically significant differences were found over time, in the interaction test, or between intervention and comparison classrooms. Mean scores for both groups of classrooms at posttest suggested that the teachers had a basic rating on this subscale, an average of a score of three

points per question with some evidence. On the Language Environment subscale, analyses indicated that there were no statistically significant differences on the time, interaction, and group tests from the repeated measures analysis. Mean scores at posttest indicated that classrooms had basic ratings or a score of three points per question with some evidence on this subscale of the ELLCO.

The analyses for the fourth and fifth subscale suggested similar findings. On the Books and Book Reading subscale, repeated measures analyses indicated there were no statistically significant differences over time, in the interaction term, or between intervention and comparison classrooms. Mean scores at posttest indicated that these classrooms had moderately strong scores of four points per question with sufficient evidence on this ELLCO subscale. On the Print and Early Writing subscale, analyses suggested that no statistically significant differences in the time analysis, interaction term, or group tests. Mean scores at posttest indicated that classrooms had below basic scores of less than three points per question with slightly more than limited evidence on this subscale of the ELLCO.

Table 6

Results of Repeated Measures Analysis of Variance for the ELLCO Measure

	Time 1 (pretest)		Time 2 (posttest)		F_T	F_{TxG}	F_G
	M	SD	M	SD	$\frac{P}{\eta^2}$	$\frac{P}{\eta^2}$	$\frac{P}{\eta^2}$
Classroom Structure					.41	.01	3.03
Intervention	14.73	2.37	15.00	2.27	.53	.94	.10
Comparison	16.33	1.63	16.67	1.75	.02	.00	.14
Curriculum					.88	1.50	.27
Intervention	9.47	2.30	9.33	2.97	.36	.27	.59
Comparison	8.33	1.03	9.33	1.37	.04	.07	.02
Language Environment					.02	.25	.77
Intervention	12.80	2.93	12.60	3.62	.90	.63	.39
Comparison	11.33	2.59	11.67	1.63	.00	.01	.04
Books, Book Reading					.86	.01	.62
Intervention	17.60	2.20	17.00	3.34	.37	.93	.44
Comparison	16.67	1.86	16.17	2.14	.04	.00	.03
Print, Early Writing					.05	.00	.32
Intervention	8.60	2.59	8.73	3.01	.82	.98	.58
Comparison	8.00	1.41	8.17	1.72	.00	.00	.02

Note: The maximum score for the ELLCO subscales are Classroom Structure (20), Curriculum (15), Language Environment (20), Book & Book Reading (25), and Print and Early Writing (15); values included for η^2 (eta squared) are .01 for a small effect, .06 for a medium effect, and .15 and higher for a large effect. *Statistical test is significant at the .05 level (2-tailed).

ELLCO scores were also examined as a function of the total number of hours of classroom support that a classroom received from coaching. To examine this question, a series of correlation analyses were conducted to examine the strength and direction of the association.

To examine the relationship of coaching hours on the intervention classrooms, a series of correlations were conducted specifically isolating the four subcategories of coaching (Public Practice, Pre and Post Conferencing, Observation, and Study Groups) to each of the five subscales of the ELLCO. This examination of subcategories is useful in determining if one subcategory of coaching is more strongly associated with the ELLCO posttest scores than others. Of the 20 correlations examined, 70% were negative (14 out of 20). Three of the coaching subcategories contained at least two negative correlations, although Observation and Study Groups contained four negative correlations each. Although several of the correlations were negative, none were significant. The correlations between the four subcategories of coaching and the five subscales of the ELLCO are found in Table 7.

Table 7

Correlation Analysis for Coaching Practices and the ELLCO Measure

	Coaching: Public Practice	Coaching: Pre and Post Conferencing	Coaching: Coach Observation	Coaching: Study Groups
	<i>r</i> <i>p</i>	<i>r</i> <i>p</i>	<i>r</i> <i>p</i>	<i>r</i> <i>p</i>
Classroom Structure Intervention	-.20 .49	-.15 .60	-.14 .62	-.26 .36
Curriculum Intervention	-.17 .55	.12 .68	-.37 .18	-.11 .69
Language Environment Intervention	.18 .52	.06 .83	-.22 .43	-.04 .87
Books, Book Reading Intervention	.09 .76	-.11 .70	.15 .58	.08 .78

	Coaching: Public Practice	Coaching: Pre and Post Conferencing	Coaching: Coach Observation	Coaching: Study Groups
	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
	<i>p</i>	<i>p</i>	<i>p</i>	<i>p</i>
Print, Early	-.29	-.05	-.13	-.12
Writing Intervention	.30	.87	.65	.67

*Statistical test is significant at the .05 level (2-tailed).

Lastly, an analysis was run on the ELLCO posttest scores for teachers who had participated in GRANT for the entire three years. An examination of these particular teachers' classrooms is useful in determining if posttest scores significantly increased from Year 1 of the longitudinal study to Year 3. Using an analysis of variance test, there were no statistically significant differences between the posttest scores of classrooms from Year 1 and Year 3. With a possible total score of 95, posttest scores from all three years ranged between 63% - 68% of a full score of 95. The secondary analysis data set for Year 3 did not include students' posttest scores or coaching hours for Years 1 and 2 so analyses for these variables was not possible. Means and standard deviations along with the specific value for the analysis of variance measure are found in Table 8.

Table 8

Results of Repeated Measures Analysis of Variance for the ELLCO Measures Over Time

			F
	<i>M</i>	<i>SD</i>	<i>p</i>
Year 1	62.27	14.60	.41
Year 2	64.91	10.68	.67
Year 3	60.27	10.61	

A graphic visual representation and description of the graph of classroom ELLCO posttest scores provides additional information to the question of the relationship between coaching hours and the ELLCO scores over time. For the 11 out of 21 teachers' classrooms, although individual classrooms increased or decreased posttest scores, analysis of variance tests found no statistically significant differences ($F = .41, p = .67$). Between Year 1 and 2, six classrooms (54%) increased their posttest scores, and five classrooms (45%) decreased their posttest scores. Between Year 2 and 3, five classrooms (45%) increased their posttest scores, and six classrooms (54%) decreased their posttest scores. Of those six classrooms that showed increases from Year 1 to 2, only two classrooms (18%) continued to increase their posttest scores from Year 2 to Year 3.

Between Year 1, 2 and 3, two classrooms (18%) consistently increased their posttest scores in all three time intervals, two classrooms demonstrated consistent decreases (18%), four classrooms (37%) demonstrated higher scores in Year 2, and three classrooms (27%)

demonstrated higher scores in Year 3. Figure 4 graphs the posttest scores for these 11 classrooms over Years 1, 2, and 3.

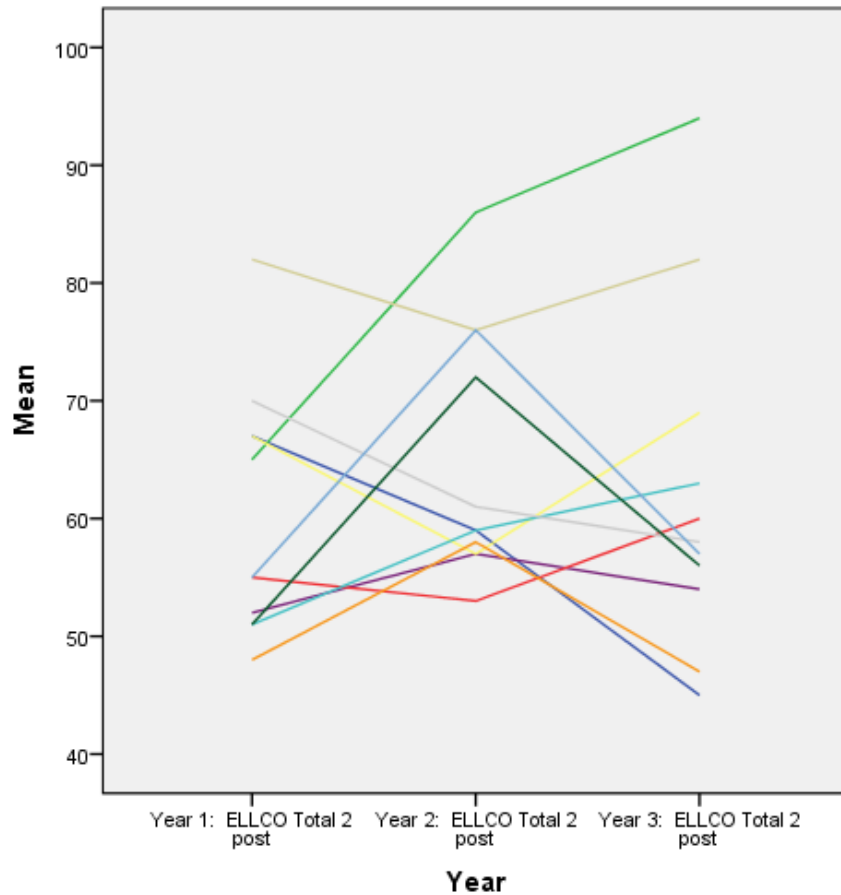


Figure 4. Graphic Representation of the ELLCO Measure Over Time.

Summary

The purpose of the study was to analyze the available secondary data from one quasi-experimental designed Early Reading First grant and understand the relationship of coaching hours on child outcomes, teacher practices, and/or changes in the classroom environment. The research questions included 1) What are the relationships between coaching hours and children’s literacy outcomes? and 2) What is the relationship between coaching hours and teacher practices and/or changes in the classroom environment? Child outcomes were measured by significant increases in the posttest scores of children’s PPVT, EVT, and PALS Pre-K assessments. Teacher

practices and classroom environments were measured by changes in the classroom posttest ELLCO scores.

In response to the first question, analysis of variance tests findings indicated that in the area of receptive language, students made significant gains pretest to posttest, but the interaction and group effects were not significant. For expressive oral language, students made statistically significant gains pretest to posttest and a significant difference was found in the group effect, indicating that students in the intervention classrooms made greater gains than students in the comparison classrooms in this particular measure. In the area of phonological awareness, students made significant gains pretest to posttest, but an interaction effect was found in the subsection of the Rhyme Awareness section only. Findings indicate positive correlations between Observation and the PPVT and EVT posttest scores. However, there were also several statistically significant negative correlations: Public Practice and Pre and Post Conferencing to the PPVT posttest scores, Pre and Post Conference to the EVT posttest scores, and Pre and Post Conference and Study Groups to the Print and Word Awareness section of the PALS Pre-K.

In response to question two, there were no statistically significant differences in the time analysis, interaction term, or group effect on any of the ELLCO subscales. On the ELLCO measure, there were no statistically significant correlations between the number of coaching hours and the ELLCO posttest scores, including teachers who participated in Years 1, 2, and 3 of the grant.

CHAPTER 5

DISCUSSION OF THE FINDINGS

The purpose of this study was to examine one Early Reading First project to shed light on the relationship of coaching hours on child outcomes and the changes that occur in teacher practices and/or the classroom environments. To add to the research and evidenced-base of the coaching literature, the following questions were addressed in this study:

- 1) What are the relationships between coaching hours and children's literacy outcomes?
- 2) What is the relationship between coaching hours and teacher practices and/or changes in the preschool classroom environment?

In 2009, Landry et al. highlighted a misalignment between the education of early childhood providers and the teaching skills needed to provide effective instruction for young children. This misalignment created the need for early childhood professional development. Suggesting that professional development in the form of coaching was one strategy with the potential to bridge the gap found in early childhood educational settings, the current study sought to add to the research base and discussion about coaching. It was anticipated that the findings of the current study might contribute to the knowledge gap in the field of early childhood on the relationship between coaching hours and child outcomes.

Summary, conclusions, and recommendations of this study are presented in this chapter, organized into separate sections. These sections include a summary of the current study's findings, an examination of the major findings related to key literature, a discussion of the findings, and the limitations to the current study. The chapter concludes with implications and

recommendations for researchers, practitioners, and policymakers, as well as the researchers' final thoughts.

Summary of the Findings

Secondary data analysis was the methodology used to address the research questions posed in this study. The analyses presented here are arranged in the order of the two research questions posed, beginning with repeated measures analysis of variance tests between children's outcome measures and coaching hours. After such analyses were conducted, children's posttest scores were correlated with coaching hours to determine if there was an effect from coaching.

The second research question addressed the extent to which the support of a literacy coach, as measured by coaching hours, was related to changes in teacher practices and/or changes in the classroom environment. Repeated measures analysis of variance was conducted to determine if additional support offered by a literacy coach affected classroom environment scores. Similarly, correlations were also computed to examine coaching hours and their relationship to environment scores. Classroom environment scores were also examined to assess teacher experience in the program and its relationship to coaching. Nine main findings of the study were noted:

Student Measures:

The Peabody Picture Vocabulary Test, Fourth Edition (PPVT):

1. Statistically significant differences were found on the time analysis for pretest and posttest student PPVT scores, but no differences were found on the interaction term or the group effect.

2. The coaching subcategories of Public Practice and Pre and Post Conference were significantly negatively correlated to the PPVT posttest scores, while Coach Observation was significantly positively correlated.

The Expressive Vocabulary Test, Second Edition (EVT):

1. Statistically significant differences were found on the time analysis for pretest and posttest student EVT scores as well as the group effect, but no differences were found on the interaction term.

2. The coaching subcategory of Pre and Post Conference was significantly negatively correlated with the EVT posttest scores, while Coach Observation was significantly positively correlated to the EVT posttest scores.

The Phonological Awareness and Literacy Screening (PALS Pre-K):

1. There were statistically significant differences on the time analysis for all four sections of the PALS Pre-K measure. There was also a statistically significant difference on the Rhyme Awareness section between posttest scores and the condition, the interaction term, but not the group effect.

2. The coaching subcategories of Pre and Post Conference and Study Groups were both statistically negatively correlated with the PALS Pre-K Print and Word Awareness section.

Teacher Measure:

The Early Language and Literacy Classroom Observation tool (ELLCO):

1. There were no statistically significant differences in the time analysis, interaction term, or group effect on any of the ELLCO subscales: Curriculum Structure, Curriculum, and Language Environment, Books and Book Reading, and Print and Early Writing subscales.

2. There were no statistically significant correlations between coaching hours and the ELLCO posttest scores. These correlations were run between total ELLCO scores and subcategories of coaching hours, including the isolation of the intervention classrooms.

3. There were no statistically significant differences in the posttest scores between Years 1, 2, and 3.

The findings of the secondary data analysis, along with GRANT's longitudinal evaluation suggest that children's pretest and posttest scores in several predetermined areas, including oral language (measured by the PPVT and the EVT assessments) and the phonological awareness skill of rhyme awareness (measured by the PALS Pre-K assessment), demonstrate statistically significant differences in their growth. Only the group effect analysis for the EVT was significant, indicating that children in the intervention classrooms showed a statistically significant difference in posttest scores compared to the children in the comparison classrooms. Only the interaction term for the Rhyme Awareness section of the PALS-Pre K was significant, meaning that students in classrooms with coaching support showed a steeper increase in posttest scores than students in classrooms without coaching support for teachers in this particular area.

In addition, there were several correlation findings indicating that coaching hours and specific measures of child outcomes were statistically significant. Findings indicated positive correlations between Observation and the PPVT and EVT posttest scores, suggesting that as Observation hours increased, so did children's PPVT and EVT posttest scores. However, there were also several statistically significant negative correlations: Public Practice and Pre and Post Conferencing to the PPVT posttest scores, Pre and Post Conference to the EVT posttest scores, and Pre and Post Conference and Study Groups to the Print and Word Awareness section of the PALS Pre-K.

In examining the relationship between coaching hours and teacher practices and/or changes in the classroom environment, there were no statistically significant differences in the time analysis, interaction term, or group effect on any of the ELLCO subscales. On the ELLCO measure, there were no statistically significant correlations between the number of coaching hours and ELLCO posttest scores, including teachers who participated in Years 1, 2, and 3 of the grant.

Major Findings Related to Key Literature and Discussion of the Findings

Beginning in the 1980s, Joyce and Showers (2002) presented the hypothesis that training components with “presentation by theory, modeling or demonstration, practice...and in-class assistance with transfer” (p. 85) were the types of training most likely to lead to various levels of child outcome and teacher performance impact. These seminal researchers posited that as teachers were expected to master new curricula, approaches of teaching, and new practices, technical assistance at the classroom level would be critical to their success.

The No Child Left Behind (NCLB) legislation in 2001 added an “intensified spotlight on teacher preparation, reading instruction, and student achievement” (Elish-Piper & L’Allier 2010, p. 162). The combination of concentrated attention on teacher preparation coupled with the concept of job-embedded professional development brought coaching to a national level of awareness as a viable practice that could influence children’s outcomes. While research on coaching is more thoroughly examined in the primary grades than in early childhood settings, there is debate regarding its effectiveness in the literature.

The coaching literature has focused on the theory of coaching and its expectations to improve teacher capacity and student achievement (Elish-Piper & L’Allier, 2011; Jackson et al., 2007; Neuman & Cunningham, 2009; Neuman & Wright, 2010), lack of evidence to support

coaching linked to student achievement and teacher capacity (Landry, 2009; Wayne et al., 2008), and the need for investigations that seek a better empirical understanding to these questions (Bean et al., 2010; Elish-Piper & L’Allier, 2010; Shidler, 2009).

Garet, Porter, Desimone, Birman, and Yoon (2001) explained that traditional forms of professional development, such as stand-alone one time workshops, were ineffective. These types of professional development activities and practices lack the time necessary for teachers to practice new skills that result in changes to practice. Coaching, with its emphasis on sustained and intense work over time, was hailed as a “reform” type of effective professional development (Garet et al., 2001, p. 920). Wayne et al. (2008) offered that the literature on coaching as a professional development strategy provided direction and offered an “informal consensus” about the features of professional development that were most effective, but the “evidence base” was weak and more specific research on professional development was necessary to “address these basic questions of policy and practice” (p. 476) related, for example, to coaching.

Other empirical research on the impacts of coaching produced varied outcomes. Specific to the literature focusing on Kindergarten through third grade, Bean et al. (2010) found that in schools in which coaches spent more time, there were a greater percentage of students scoring proficiently on one measure of achievement in first and second grades. Elish-Piper and L’Allier (2010, 2011) suggested that time, measured by a coach log, may be related to student reading achievement. The present study focused specifically on the variable of coaching hours in preschool classrooms, thus providing limited support to the previous two studies. The present findings demonstrated only the posttest expressive vocabulary measure showed statistically significant differences in the group effect, suggesting that the effect of coaching support for the

children in the intervention classrooms made a statistically significant difference in their relative performance on the posttest EVT measure.

The current study's analysis of multiple measures of child assessment provided more support to Garet et al.'s (2008) study. Commissioned by the Institute of Education Sciences (IES) and the National Center for Education Evaluation and Regional Assistance (NCEE), Garet et al.'s study found little effects of coaching to classroom instruction and student achievement. The differences in the focal variables, design of the study, and age of the students made comparisons between these studies and the current study complex and although the studies are functionally different, all add to the literature of evidenced-base coaching practices.

Specific to the field of early childhood education, the current study is most related to three studies, although various differences still exist. Shidler (2009) reported significant correlation between hours of coaching and letter recognition, and no significant correlation between the ranked classroom PPVT scores and coaching hours. On a measure of letter recognition, the findings from the current study are inconsistent with these results, finding no statistically significant difference in posttest scores between intervention and comparison classrooms on Upper Case letter recognition, but are consistent with the PPVT findings from Shidler's (2009) study.

Neuman and Cunningham (2009) were able to isolate the combination of variables of coursework and coaching. The combination of these two variables produced statistically significant effects on improvements in teacher practices, suggesting that coursework alone produced "negligible effects on improvements in quality practices" (p. 532). Neuman and Wright (2010) build on earlier research to address the gaps they previously identified. With a research group that isolated the coaching variable from the effects of coursework alone, they

found that the coaching group evidenced a statistically significant effect on teacher practices, as measured by the ELLCO. The findings from the current study stand in contrast to Neuman and Wright's (2010) analysis in that there was no correlation between changes in teacher practices, changes in the classroom environment, and hours spent coaching.

In 2007, the Department of Education (DOE) and the Institute of Education Sciences (IES) released its evaluation and findings of ERF programs (Jackson et al., 2007). From the longitudinal evaluation and secondary data analysis, the current study's results run counter to Jackson et al.'s time analysis results. The present findings indicate statistically significant differences in the time analysis in all child measures, both receptive and expressive language, as well as the four areas of phonological awareness. These findings also add to the research base by reporting a statistically significant group effect for the expressive measure, suggesting that the effect of coaching support for intervention classrooms made a statistically significant difference in their relative performance on the EVT posttest measure.

The current study also added to the literature by addressing the variable of coaching hours and its relationship to teacher practices and/or changes in the preschool classroom. Although none of the correlations are statistically significant, 70% of the relationships were negative, suggesting an inverse relationship between coaching hours and teacher practices. These mixed results on multiple correlation analyses add to the debate on the relationship of coaching hours on child outcomes. The often conflicting correlation results could be very contextual and add credence to the call for more research in early childhood settings specifically looking at the relationship of coaching hours to child outcomes, teacher practices, and/or changes in the classroom environment. The research implications and additional questions posed are discussed in the following section.

Limitations to the Current Study

GRANT was quasi-experimental in design, and involved the very real and practical scenarios of teaching; therefore, it focused primarily on implementation. For the current study, the data set used consisted of secondary data collected by an outside evaluator. The researcher was granted permission to analyze the data that were collected, but did not have any influence over the original evaluation design. As a result of missing and/or incomplete data, the current study was unable to make strong connections between specific variables in the intervention classrooms, but was able to offer general conclusions based on the findings. In GRANT'S original design and in its authentic teaching environment, there were factors for which the study did not attempt to account, such as teacher motivation, parental engagement, and natural child maturation. As a result, there was difficulty in identifying the variable of coaching hours as the sole or even primary reason for the increased student outcomes.

From a research perspective, the study was conducted as a quasi-experimental design with four coaches working with the intervention group and a small number of teachers and children in the comparison group. Samples including larger numbers of participants could yield stronger and more comparable results.

Implications and Recommendations for Researchers

Based on these findings, the following section outlines implications and recommendations for researchers, practitioners, and policymakers. It is important to situate these implications and recommendations within the context on the study's findings and specific geographic setting. The current study's questions and findings imply there is merit in isolating specific variables within the strategy of coaching as professional development. The study also

offers additional questions to the field of early education and the on-going research of effective professional development strategies, namely coaching.

The research and data findings presented in this study were designed to shed light on the relationship of coaching hours on children's literacy outcomes, teacher practices, and changes in the preschool classroom environment. This line of inquiry also holds interest for researchers studying financial investments in early education at the state and federal levels, specifically those allocated through grants and other competitive initiatives.

Early Reading First was designed to "transform existing early education programs into centers of excellence that provide high-quality, early education to young children, especially those from low-income families" (U.S. Department of Education, 2012, ¶ 3). In the case of the local non-profit organization (LNP) in the current study, the number of coaching hours dedicated to working with ERF teachers indicates that LNP sought to increase the quality of education for children by strengthening teacher effectiveness; therefore, increasing child outcomes with these posttest scores as the final unit of measurement.

On their own research in coaching and student achievement in Kindergarten through third grade, Elish-Piper and L'Allier (2011) recommended that coaching logs be further developed and refined in their methods and tools used to collect data about how literacy coaches spent their time, how they engaged teachers, and around what content areas. Based on the current study, researchers should identify and address specific types of dosage and frequency questions to narrow the focus on coaching variables and their relationship on teacher practices and student achievement. With a greater pool of comparable data across projects in various locations and populations, researchers could offer more specific direction as to the most cost effective and productive types of coaching activities.

Implications and Recommendations for Practitioners

The implications to practitioners include authentic opportunities to consider carefully the associations and possible anecdotal roots of these relationships and associations. If coaching hours are influencing or seem to indicate associations with student achievement, is the time spent with coaches the true cause? Perhaps more focused planning time with other teachers or specific attention to one teaching strategy would create similar results. Although the study's correlation findings imply an association, or more specifically, the lack of association, school leadership would be interested in practically addressing the question of identifying more specifically the factors of effective professional development. If student posttest scores increased significantly, but coaching hours were not a major contributing factor, what were the other factors? How can the use or misuse of a coaching model influence its relationship to student achievement? What coaching qualifications, educational, and interpersonal skills, are most helpful in achieving the desired results and how do the levels of experience and receptiveness of the teaching staff fit into these overall decisions? These conversations include seriously identifying and vigorously studying all the possible input factors associated with professional development and the coaching strategy. The close examination of a professional development system could provide not only more supportive academic systems, but also drive more sound financial decisions, as well.

Based on the current study's findings, recommendations for school leaders include encouraging them to collect and analyze their own data with similar questions in mind. What forms of professional development and/or types of coaching result in increased student achievement and changes in classroom environments? While this answer would be helpful, extending the question beyond the identified form or type of professional development offers additional valuable information. Practitioners could use their initial findings and consider the

dosage and frequency of a specific activity. To increase the usefulness of their findings, practitioners should also consider ways to validate their internal data collection process. For example, teachers could videotape lessons for peer review, and coaches might consider independent observations to maintain consistent coding practices.

From a school budgetary prospective, what types of human resources and financial investments are necessary to see changes in teacher practices, how long before the changes are sustainable, and in what environments are the changes likely to become internalized? These questions address the capacity and sustainability of any chosen form of professional development and should be examined as part of both preplanning activities and ending evaluations. Even at a classroom, grade, or building level, this line of questioning and the resulting development of unique hypotheses are worth testing if significant student achievement and increased teacher knowledge and capacity are possible.

Implications and Recommendations for Policymakers

Implications for policymakers include recognizing the effects of professional development on a larger scale and on specific programs, such as grants and other federal funding. Temple and Reynolds (2007) offered that in recent years, estimates suggesting a high rate of return to early childhood education have more than piqued the interest of policymakers. The policies supporting and program opportunities calling for an investment in youth and children are immense. However, additional questions involving the length of investment of an intervention, in this case participation in a grant, are monetarily relevant. During a multi-year grant, what progress can reasonably be expected, when, and in what order? Are changes in teacher practices indicators of increased instructional knowledge or is the reverse more logical? Does it matter if student scores are improving? Under what conditions are these evidence-based practices and

successes scalable, replicable, and/or sustainable? Heckman (1999) explained that “in evaluating a human capital investment strategy, it is crucial to consider the entire policy portfolio of interventions together- training programs, school-based polices, school reform, and early intervention- rather than focusing on one type of policy in isolation from the others” (p. 138). Policymakers could consider these questions to aid in framing and then in making effective financial investment decisions in education.

Considering the policy arena, data from this study suggest that federal grant discussions and planning decisions carefully consider the questions they are interested in and collect data most useful in addressing those questions. With an awareness of the need for both fiscal and long term returns, policymakers might consider working with researchers and practitioners to craft, assess, and modify these large scale programs. While all educational settings are unique because of the children they serve, some similarities still exist. Policymakers might consider what aspects of a multi-year grant would be beneficial to standardize to increase the pool of reliable and useful data. Secondary data analysis is cost beneficial, but only if comparable information is available.

Final Thoughts

The goal of any scholarly work is to arguably add to the research and/or evidence base of a particular topic. In the case of coaching as an effective professional development strategy, researchers and practitioners are wise to remain open to the broader questions their research presents, for fear of focusing too narrowly on less relevant questions. The issue is larger than a simple binary answer, too complex an inquiry for a simple yes or no response. Without overextending its findings, the significance of this study is perhaps not in the answers it provides, but in the questions it encourages and challenges the field to address. As Pierre Marc Gaston de

Lévis, Duke of Lévis (1764-1830), wisely offered, “Judge a man by his questions rather than his answers.”

References

- Barnett, S. (2012). *The impact of quality ECEC on child development: Designing effective policy and practice*. Unpublished manuscript, National Institute for Early Education Research, New York.
- Bean, R. M., Draper, J. A., Hall, V., Vandermolen, J., & Zigmond, N. (2010). Coaches and coaching in reading first schools. *Elementary School Journal, 111*(1), 87-114. doi: 10.1086/653471
- Brown, C., Stroh, H., Fouts, J., & Baker, D. (2005). *Learning to change: School coaching for systemic reform*. Seattle, WA: Bill and Melinda Gates Foundation.
- Campbell, J. M., Bell, S. K., & Keith, L. K. (2001). Concurrent validity of the peabody picture vocabulary test—third edition as an intelligence and achievement screener for low ses african american children. *Assessment, 8*(1), 85-94. doi: 10.1177/107319110100800108
- DeBray-Pelot, E., & McGuinn, P. (2009). The new politics of education: Analyzing the federal education policy landscape in the post-NCLB era. *Educational Policy, 23*(1), 15-42. doi: 10.1177/0895904808328524
- DellaMattera, J. (2010). No preschooler left behind: Preschool policies in the NCLB World. *Journal of Educational Research & Policy Studies, 10*(1), 35-49.
- Deussen, T., Coskie, T., Robinson, L., & Autio, E. (2007). “Coach” can mean many things: Five categories of literacy coaches in Reading First (Issues & Answers, REL 2007-No. 005). Retrieved December 30, 2012 from U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, <http://ies.ed.gov/ncee/edlabs/projects/project.asp?ProjectID=47>.

- Dickinson, D., & Caswell, L. (2007). Building support for language and early literacy in preschool classrooms through in-service professional development: Effects of the Literacy ENVIRONMENT Enrichment Program (LEEP). *Early Childhood Research Quarterly*, 22, 243-260. doi: 10.1016/j.ecresq.2007.03.001
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). Bloomington, IN: Pearson Assessments.
- Elish-Piper, L., & L'Allier, S. K. (2010). Exploring the relationship between literacy coaching and student reading achievement in grades k-1. *Literacy Research & Instruction*, 49(2), 162-174. doi: 10.1080/19388070902913289
- Elish-Piper, L., & L'Allier, S. K. (2011). Examining the relationship between literacy coaching and student reading gains in grades k-3. *Elementary School Journal*, 112(1), 83-106. doi: 10.1086/660685
- Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., . . . Silverberg, M. (2008). *The impact of two professional development interventions on early reading instruction and achievement*. Washington, DC: Institute of Education Sciences [and] National Center for Education Evaluation and Regional Assistance.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915-945. doi: 10.3102/0002831207308221
- Heckman, J. J. (1999). *Policies to foster human capital* (working paper 7288). Cambridge, MA: National Bureau of Economic Research.

- Holland, P. (2005). The case for expanding standards for teacher evaluation to include an instructional supervision perspective. *Journal of Personnel Evaluation in Education*, 18(1), 67-77. doi: 10.1007/s11092-006-9009-0
- Invernizzi, M., Sullivan, A., Meier, J., & Swank, L. (2004). *Phonological awareness and literacy screening (pals) prek technical manual*. Richmond, VA: University of Virginia.
- Jackson, R., McCoy, A., Pistorino, C., Wilkinson, A., Burghardt, J., Clark, M., . . . Schmidt, S. (2007). *National evaluation of early reading first: Final report to congress*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.
- Joyce, B., & Showers, B. (2002). *Student achievement through staff development*. New York, NY: Longman.
- Kretlow, A. G., & Bartholomew, C. C. (2010). Using coaching to improve the fidelity of evidence-based practices: A review of studies. *Teacher Education & Special Education*, 33(4), 279-299. doi: 10.1177/0888406410371643
- Landry, S. H., Anthony, J. L., Swank, P. R., & Monseque-Bailey, P. (2009). Effectiveness of comprehensive professional development for teachers of at-risk preschoolers. *Journal of Educational Psychology*, 101(2), 448-465. doi: 10.1037/a0013842
- Landry, S. H., Swank, P. R., Smith, K. E., Assel, M. A., & Gunnewig, S. B. (2006). Enhancing early literacy skills for preschool children: Bringing a professional development model to scale. *Journal of Learning Disabilities*, 39(4), 306-324. doi: 10.1177/00222194060390040501
- Local Non-Profit (LNP) (2009). [LNP Narrative_Education].

- Lockwood, J. R., Sloan McCombs, J., & Marsh, J. (2010). Linking reading coaches and student achievement: Evidence from florida middle schools. *Educational Evaluation & Policy Analysis*, 32(3), 372-388. doi: 10.3102/0162373710373388
- McCrosky, H. (2010). [Codes for coaching logs].
- McGreal, T. (1997). Can a supervisor be a coach? In J. Glanz & R. Neville (Eds.), *Education supervision: Perspective, issues, and controversies* (pp. 91-99). Norwood, MA: Christopher-Gordon Publishers, Inc.
- National Association for the Education of Young Children & National Association of Child Care Resource & Referral Agencies. (2011). *Early childhood education professional development: Training and technical assistance glossary*. Retrived May 12, 2012 from http://www.naeyc.org/GlossaryTraining_TA.pdf
- Neuman, S. B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal*, 46(2), 532-566. doi: 10.3102/0002831208328088
- Neuman, S. B., & Wright, T. S. (2010). Promoting language and literacy development for early childhood educators. A mixed-methods study of coursework and coaching. *Elementary School Journal*, 111(1), 63-86. doi: 10.1086/653470
- Nolan, J. F. (1997). Can a supervisor be a coach? In J. Glanz & R. Neville (Eds.), *Education supervision: Perspective, issues, and controversies* (pp. 100-108). Norwood, MA: Christopher-Gordon Publishers, Inc.
- Onchwari, G., & Keengwe, J. (2008). The impact of a mentor-coaching model on teacher professional development. *Early Childhood Education Journal*, 36(1), 19-24. doi: 10.1007/s10643-007-0233-0

- Powell, D. R., Diamond, K. E., Burchinal, M. R., & Koehler, M. J. (2010). Effects of an early literacy professional development intervention on head start teachers and children. *Journal of Educational Psychology, 102*(2), 299-312. doi: 10.1037/a0017763
- Restrepo, M. A., Schwanenflugel, P. J., Blake, J., Neuharth-Pritchett, S., Cramer, S. E., & Ruston, H. P. (2006). Performance on the ppvt-iii and the evt: Applicability of the measures with african american and european american preschool children. *Language, Speech, and Hearing Services in Schools, 37*(1), 17-27. doi: 10.1044/0161-1461(2006/003)
- Rogers, W. T., Anderson, J. O., Klinger, D. A., & Dawber, T. (2006). Pitfalls and potential of secondary data analysis of the council of ministers of education, Canada, national assessment. *Canadian Journal of Education/Revue Canadienne de l'éducation, 29*(3), 757-770. doi:10.2307/20054194
- Shanklin, N. (2006). What are the effective characteristics of effective literacy coaching? Literacy Coaching Clearinghouse. Retrieved May 2, 2012, from <http://www.usdb.org/curr/literacy/Administrator%20Materials/Characteristics%20of%20Literacy%20Coaching.pdf>
- Shanklin, N. (2009). Literacy coaching: What are we learning? CEDER Yearbook, 31-44, Corpus Christi, TX: Texas A&M University-Corpus Christi, Center for Educational Development, Evaluation & Research.
- Shidler, L. (2009). The impact of time spent coaching for teacher efficacy on student achievement. *Early Childhood Education Journal, 36*(5), 453-460. doi: 10.1007/s10643-008-0298-4

- Smith, M. W., Brady, J. P., & Anastasopoulos, L. (2008). *User's guide to the early language and literacy classroom observation pre-k tool*. Baltimore, MD: Paul H. Brooks Publishing Co.
- Smith, S., Robbins, T., Schneider, W., Kreader J. L., & Ong, C. (2012). *Coaching and quality assistance in quality rating improvement systems: Approaches used by ta providers to improve quality in early care and education programs and home-based settings*. New York: National Center for Children in Poverty.
- Sturtevant, E. G. (2003). *The literacy coach: A key to improving teaching and learning in secondary schools*. Washington, DC: The Alliance for Excellent Education.
- St. Cyr, J. (2012). *GRANT: Annual evaluation report*. Atlanta, GA: The Evaluation Group.
- Temple, J. A., & Reynolds, A. J. (2007). Benefits and costs of investments in preschool education: Evidence from the child–parent centers and related programs. *Economics of Education Review*, 26, 126-144. doi: 10.1016/j.econedurev.2005.11.004
- U.S. Department of Education. (2012). *Archived information, Purpose: Early reading first*. Retrieved June 10, 2012, from <http://www2.ed.gov/programs/earlyreading/index.html>
- U.S. Department of Education. (2009a). *Archived information, Performance: Early reading first*. Retrived June 12, 2012 from <http://www2.ed.gov/programs/earlyreading/performance.html>
- U.S. Department of Education. (2009b). *Archived Information, Awards: Project abstracts FY 2009 early reading first grant program*. Retrived April 16, 2012 from <http://www2.ed.gov/programs/earlyreading/awards.html>
- U.S. Department of Education. (2003). *Archived Information, Guidance for the early reading first program*. Retrived March 16, 2012 from <http://www2.ed.gov/programs/earlyreading/legislation.html>

- Vanderburg, M., & Stephens, D. (2010). The impact of literacy coaches. *Elementary School Journal*, 111(1), 141-163. doi: 10.1086/653473
- Wayne, A., Yoon, K., Zhu, P., Cronen, S., & Garet, M.S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher*, 37(8), 469-479. doi: 10.3102/0013189X08327154
- Walpole, S., McKenna, M., Uribe-Zarain, X., & Lamitina, D. (2010). The relationship between coaching and instruction in the primary grades. *The Elementary School Journal*, 3(1), 115-140. doi: 10.1086/653472
- Williams, K. T. (2007). *Expressive vocabulary test* (2nd ed.). Bloomington, IN: Pearson Assessments.
- Zepeda, S. (2012a). *Professional development: What works* (2nd edition). Larchmont, NY: Eye On Education, Inc.
- Zepeda, S. (2012b). *Job-embedded learning: A powerful, practical, and cost-effective form of professional development* [White paper]. Retrieved from http://eyeoneducation.com/bookstore/client/client_pages/pdfs/Zepeda_Job_Embedded_Learning.pdf