

Copyright
by
Christa Dawn Haring
2013

The Dissertation Committee for Christa Dawn Haring

certifies that this is the approved version of the following dissertation:

**The Effects of Coaching on Teacher Knowledge, Teacher Practice and Reading
Achievement of At-Risk First Grade Students**

Committee:

Sharon Vaughn, Supervisor

Carolyn Denton

Andrea Flower

Ed Emmer

Nina Zuna

The Effects of Coaching on Teacher Knowledge, Teacher Practice and Reading

Achievement of At-Risk First Grade Students

By

Christa Dawn Haring, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2013

**The Effects of Coaching on Teacher Knowledge, Teacher Practice and Reading
Achievement of At-Risk First Grade Students**

Christa Dawn Haring, Ph.D.

The University of Texas at Austin, 2013

Supervisor: Sharon Vaughn

Dedication

For my parents

Acknowledgements

Reinhold Niebuhr taught “nothing we do, however virtuous, can be accomplished alone.” Completing my doctoral program is evidence of this. I could easily fill as many pages with the names of those who have contributed to the completion of this process as I have with the text describing the study itself. More than an academic endeavor, education, in all its forms, is less about completing and more about becoming. It is my deepest desire as I continue in my career to become a mirror – a reflection of the encouragement, kindness, wisdom and character of those that have helped shape who I am as a student, teacher, researcher and friend.

To Drs. Howard Goldstein and Juliann Woods, who encouraged me to begin the journey and followed my progression, I am truly grateful. My sincere thanks as well go to Dr. Don Deshler who mentored me academically and spiritually during the journey and whom I hope to better emulate as I move forward. To Dr. Jim Knight, who taught me the true meaning of ‘coaching’ through example. To Dr. Sharon Vaughn, who took a chance on an unknown applicant, taught her how to make an impact in the field and funded all of her crazy ambitions, thank you very much. To Dr. Diane Bryant, who looked into the faces of a group of incoming doctoral students and told us we could succeed. Additionally, I would like to thank Drs. Ed Emmer and Nina Zuna for their comments, suggestions and time serving on my committee; each went above and beyond on my behalf. Last but not least, my thanks go to Dr. Andrea Flower, who showed me it was possible to balance all of the important things in life and who not only served as an incredible mentor, but also became a friend.

After finishing my coursework at UT Austin, I was privileged to add to my extended family an amazing group of people from Children’s Learning Institute at UT Houston. I will forever be grateful for Dr. Carolyn Denton who provided the platform on which I could complete my dissertation and helped me find my confidence as a researcher. Her friendship and mentorship have left a great impression. Additionally, I owe a great deal to Becky Beegle and Kelli Jimenez who became my “partners in crime” and cheered with me as I grew through the process. Thanks go to Anita Najhawan, Pat Brinson, Lisa Pruitt, Elizabeth Modad and Nancy Vanderlinden who echoed their cheers and provided further examples to follow.

Along the way, I was privileged to make some incredible friends, each of whom helped shape who I am today. Many heartfelt thanks go to Lisa Belliston – who showed me it could be done. To John McKenna and Sabrina Lo, who walked the same path and helped shoulder the load. To Kim Shumake, Miss Gwen, Stephanie Townsend, Julie Lamb, Stefanie Roemer, Kathy Peterson, Austin Swafford, Whitney Smith and Melanie Pouwer, who made me smile along the way. To the Rundquist family, who

adopted me while I was away from home and treated me as one of their own, I am so very grateful. To Brenda Fields and her boys, for a listening ear and a very special jar of rocks. To Audrei Moffit Powelson, for Christmas trees on top of cars, late night talks and singing certain TV theme songs at the top of our lungs. To Mirza Lugo-Neris and Barel Alcantara, we've come a long way since graduate school but I'm pretty sure I will always need a timer. Thanks for teaching me to laugh through the tough parts. To Sam Haring, who reminded me I don't have to know everything about statistics. Finally, to Kazuyoshi Hayashi, for studying turtles, eating Frostys, buying APA manuals, and for sharing each step of the way – I will be eternally grateful.

Above all, my gratitude goes to my family. I am deeply indebted and thankful for your constant support, encouragement, examples, and testimonies. To Leah, for never letting me settle for less than I could be. For Sterling (and Julia), for showing me what you can accomplish when you follow your dreams. For Fuller (and Mel), for fighting so I had the freedom to do so and for sharing your monkeys with me.

For my Dad, for singing into the intercom to wake us up, early morning seminary, leaf collections, innumerable spiritual discussions while kayaking and for “choosing” us. May we always “remember who we are.” For my Mom, for early morning breakfasts discussing my future house across the street, for home décor inspiration found in the most unusual places, for raising me to be a part of a “long line of strong women” and for a friendship that began before our “turn on earth.” May there always be “no empty chairs.”

**The Effects of Coaching on Teacher Knowledge, Teacher Practice and Reading
Achievement of At-Risk First Grade Students**

Christa Dawn Haring, Ph.D.

The University of Texas at Austin, 2013

Supervisor: Sharon Vaughn

Abstract: The effects of coaching on teacher and student outcomes were compared to outcomes of classes randomized to professional development only and comparison conditions. Twenty-one teachers, trained to implement a Tier II reading intervention curriculum, were grouped by campus then randomized to one of three conditions: professional development plus coaching support (n=6), professional development only (n=7) and a comparison condition (n=8). Teachers in the coached and professional development only (un-coached) conditions were compared on measures of teacher knowledge and implementation fidelity as an indication of teacher practice. Student achievement scores on word attack, reading fluency and reading comprehension measures were compared for students in each of the three conditions.

A multiple-gating procedure was used to help teachers identify the five lowest-performing readers in their first grade classrooms. Students completed a battery of seven

reading ability assessments prior to and immediately following teacher-implementation of a seventeen-week reading intervention curriculum. Results of ANCOVA analyses indicated students in classes of teachers who received professional development and coaching support did not demonstrate significantly higher scores than teachers who only received professional development training on a battery of reading measures. Further analysis indicated students in the professional development plus coaching condition did have significantly higher scores than those in comparison conditions on five of the seven outcome measures.

Teachers in coached and professional development only conditions completed a teacher knowledge survey to measure their knowledge of evidence-based reading practices. ANCOVA analysis revealed no significant differences between groups at posttest. Changes in teacher practice were measured as a function of intervention implementation fidelity. Intervention teachers were videotaped three times over the course of the intervention and taped classes were scored, rated and compared across conditions. Results of a Repeated Measures Analysis of Variance on fidelity scores revealed a statistically significant difference in favor of the teachers who received professional development plus coaching.

Table of Contents

List of Tables	xii
List of Figures	xiii
Chapter 1: Introduction.....	1
Statement of Purpose and Research Questions	10
Chapter 2: Review of Literature.....	13
Evidence Based Reading Instruction	13
Teacher Knowledge and Preparation.....	15
Staff Development	16
Existing Research on Coaching.....	17
Chapter 3: Method.....	32
Overview	32
Research Questions.....	33
Research Design	36
Participants	39
Setting	41
Participant Selection	42
Procedure and Conditions	43
Interventions.....	44
Training.....	50
Measures	51
Data Analysis	55

Chapter 4: Results	57
Data Analysis	60
Teacher Outcomes	61
Student Outcomes	67
Chapter 5: Discussion.....	83
Teacher Knowledge	84
Teacher Practice	86
Student Reading Achievement	88
Summary and Analysis	90
Future Research Recommendations	95
Limitations	96
Appendix A Reading Rules Coaching Fidelity Form	99
Appendix B Reading Rules Fidelity and Quality of Implementation Instrument	100
Appendix C Reading Rules Teacher Knowledge Survey.....	109
References.....	113
Vita	126

List of Tables

Table 1 Teacher Demographic Information.....	39
Table 2 Student Demographic Information.....	40
Table 3 Outcome Measures and Analysis.....	56
Table 4 Teacher Knowledge Survey Pretest, Posttest and Adjusted Means.....	61
Table 5 Teacher Knowledge Survey: Analysis of Covariance Results.....	62
Table 6 Teacher Implementation Fidelity.....	64
Table 7 ANOVA Results: Between Teachers Effects.....	65
Table 8 ANOVA Results: Within Teachers Effects.....	65
Table 9 Pre-test-Posttest Correlations for Student Outcome Measures.....	66
Table 10 Student Word Reading Outcomes: Homogeneity of Variances.....	67
Table 11 Word Reading: Pretest, Posttest and Adjusted Posttest Mean Scores.....	68
Table 12 Post-hoc Pairwise Comparisons for Adjusted Posttest Means.....	70
Table 13 Results of Analysis of Covariance for Word Reading Measures.....	70
Table 14 Student Outcomes: Pretest to Posttest Adjusted Mean Difference Effect Sizes.....	72
Table 15 Reading Fluency: Homogeneity of Variances.....	74
Table 16 Reading Fluency: Pretest, Posttest and Adjusted Posttest Mean Scores.....	75
Table 17 Results of Analysis of Covariance: Reading Fluency Measures.....	76
Table 18 Reading Fluency Outcomes: Post-hoc Pairwise Comparisons	77
Table 19 Reading Fluency Outcomes: Adjusted Mean Difference Effect Sizes.....	77
Table 20 Reading Fluency: Pretest, Posttest and Adjusted Posttest Mean Scores.....	78
Table 21 Results of Analysis of Covariance: Reading Comprehension.....	79

Table 22 Student Reading Comprehension Outcomes: Post-hoc Pairwise Comparisons for Adjusted Means.....	79
Table 23 Reading Comprehension Outcomes: Adjusted Mean Difference Effect Sizes.....	80

List of Figures

Figure 1 Study Design.....	38
Figure 2 Reported Coaching Activities.....	45

CHAPTER 1: Introduction

Statement of the Problem

Early performance in areas such as phonemic awareness and phonics are strong predictors of later reading fluency and comprehension abilities (McNamara, Scissons, & Gutknecht, 2011), skills considered vital to appropriately function in society (Lonigan & Shanahan, 2008). Mastery of essential literacy skills (i.e. phonological awareness, phonics, fluency, comprehension and vocabulary) is highly predictive of later educational success. Low-achieving students exhibit lower motivation for learning, decreased school engagement, and have an increased likelihood of grade retention and recidivism; difficulty with reading, specifically, is cited as the primary predictor of these poor life outcomes (Reschly, 2010).

Currently, two-thirds of fourth grade students cannot read grade-level text proficiently (NCES, 2011). Extant research documents the practices that have been proven to increase the rate of reading skill acquisition in struggling readers, yet many at-risk readers are placed in classrooms with teachers ill-prepared to implement these practices (Haycock, 2011). Sufficient training in evidence-based practices paired with sustained implementation support is requisite if we are to strengthen instruction and close the achievement gap between at-risk readers and their typically developing peers.

Readers At Risk

While national reading scores are indiscriminately low, risk factors such as low-income, minority status, and limited English proficiency significantly increase the likelihood that a child will not develop reading abilities with adequate rate or proficiency. Currently, only eighteen percent of children who qualify for free or reduced lunch are proficient readers by fourth grade (Aud, et al., 2011; NCES, 2011). Ethnically diverse students constitute forty-five percent of the total school population and perform an average of twenty-six percent lower than non-minority peers on national reading assessments, a gap that has persisted with little fluctuation in recent years (NCES, 2011). One quarter of students in high poverty schools also exhibit limited English proficiency, and twenty-one percent of these students speak a language other than English in the home; both factors that serve as potential barriers to adequate rates of reading growth (Aud, et al., 2011; Rathburn & West, 2004).

As the number of risk factors (i.e. poverty, low parent education, and home languages other than English) increase, the slope of growth in reading and mathematic skill acquisition decreases accordingly (Rathburn & West, 2004). Students entering school with little prior access to print materials lag behind peers from more academically-enriched environments, a gap that tends to remain or widen as they progress through early school years (Haycock, 2011). Describing what is commonly known as the “Mathew Effect,” McNamara, Scissons, and Gutknecht (2011) found “as children progressed from kindergarten to grade 3, those in lower ranks of reading achievement

were likely to remain in the lower ranks and furthermore, at each progressing data collection point, struggling readers fell behind their grade-level peers” (p. 421).

Teaching At-Risk Readers

Explicit, systematic instruction and an increased use of evidence-based practices have been found effective in decreasing skill deficits and improving outcomes for struggling readers (Swanson & Hoskyn, 2001; Vaughn, Gersten, & Chard, 2000). These students require intensive small group or one-on-one instruction, greater access to and interaction with print materials, more opportunities to engage with text, appropriate language models and increased opportunities to practice language and literacy skills to make sufficient growth (Ambruster, Lehr & Osborn, 2011). While proven, these practices are infrequently and inadequately implemented (Vaughn, Gersten & Chard, 2000). The recent restructuring of special education support, including the response-to-intervention framework, resulted in increased numbers of at-risk students in general education classrooms. This “blurring of special education,” referred to by Fuchs, Fuchs and Stecker (2010), often places the most at-risk students with the least qualified instructors (Haycock, 2011).

Teacher Preparation and Performance

The positive correlation between teacher knowledge and student achievement has been well-documented (Lyon & Weiser, 2009; Mather, Bos, & Baburp, 2001; Moats,

2004). Students with effective instructors make more than double the academic gains of peers in classrooms of less-experienced educators (Haycock, 1998). Further, students who retain proficient educators over consecutive years demonstrate markedly higher growth than peers in typical classrooms (Haycock, 1998).

Teacher knowledge of reading development is a significant predictor of student reading improvement (Lane, et al., 2009; Lyon, 2009; Moats, 2004). Yet, a study by Lyon (2009) found most entry-level educators have only completed two required courses in reading instruction prior to entering the classroom. Mather, Bos & Baburp (2001), in their study on educator preparation, reported that a majority of teachers' understanding and application of language was markedly inadequate to teach students with reading difficulties.

Torgesen (2000) suggests that half of our hardest-to-teach readers can acquire on-level reading ability with appropriate and early intervention, but such gains require significant change in both teacher preparation and practice. The fact that teacher knowledge is intricately connected to students' academic growth (Rupley, 2011), combined with the lack of preparation the evidence suggests teachers receive prior to entering the classroom, highlights a second performance gap that needs to be closed—that of the educators themselves.

Professional Development

The No Child Left Behind (NCLB) legislation of 2008 mandated the use of evidence-based instructional practices in the classroom and brought research-proven teaching techniques to the forefront of school reform movements. Aiming to eliminate disparate student achievement scores and publish research-proven instructional methods, institutions of higher learning have been charged with identifying effective educational practices and educating pre-service teachers in their implementation. Likewise, professional developers face a similar challenge with teachers already in the field.

While an important avenue for relaying information, professional development workshops and trainings alone have been found to be an insufficient vehicle for creating lasting classroom change (Lane, et al, 2009). Historically, professional development has taken the form of workshops, short teacher development courses, and after-school booster sessions. School districts report more than one-third of their annual operating budgets are spent on trainings and other efforts to support instructional improvement (Birman, et al., 2007). It is estimated that ninety percent of teachers nationwide attend one or more hours of professional development in reading instruction during a typical school year, yet only ten to fifteen percent of these teachers implement the strategies and even fewer continue to incorporate skills six months after training (Birham, et al., 2007; Showers & Showers, 2002).

One reason typical professional development trainings have been found ineffective is their failure to provide sufficient opportunities for the learning and practice of new skills (Cohen and Ball, 1999). Few trainings provide the adequate duration and intensity of support requisite for creating lasting change in teacher practice (Garet, Porter, Desimone, Birman, & Yoon, 2001; Boyle, While & Boyle, 2004). Research indicates these one-time trainings may increase the presence of evidence-based practices in the classroom, but that effective implementation requires a level of on-going coaching support. (Kretlow & Barlow, 2010; Kretlow, Wood, & Cooke, 2009).

Coaching

Coaching - sustained professional development via ongoing observations, planning, co-teaching, modeling and feedback by a trained professional – is not novel. As early as 1981, Joyce and Showers described a system of “support, companionship and assistance” provided by colleagues that later came to be known as peer coaching (Ackland, 1991). In the 1990s, the cognitive coaching movement shifted the emphasis from peer evaluation to a more self-reflective process with a goal of increasing teacher autonomy (Garmston, Linder, & Whitaker, 1993). Most recently, Knight’s instructional coaching model incorporates a four-pronged approach providing support in content planning, assessment, classroom management and instruction. Using combined practices of observation, feedback and support to strengthen educator practice, instructional coaches aim to increase the number of research-supported teaching techniques used in typical practice (Knight, 2007, Knight & Cornett, n.d.).

Government initiatives such as Reading First and Early Reading First provided funding for site-based literacy coaches for many preschool and elementary campuses and are credited with the rapid spread of coaching practices across the country. In 2001, in an effort to have all students reading at grade level, the Just Read Florida! initiative provided funds for more than 2,300 coaches for schools failing to meet adequate yearly progress (Marsh, Lockwood & McCombs, 2010). Although results have been mixed, positive student achievement outcomes and high social validity ratings have spurred a rapid increase in the number of coaches employed in districts across the country (International Reading Association, 2004).

In his seminal study of the additive effects of coaching support, Bush (1984) collected data on eighty teachers in twenty school districts over a five-year period. He found that adding modeling, practice and feedback to an initial professional development training increased the likelihood of implementation by two to three percent per technique, resulting in implementation rates of only eleven to sixteen percent of these new practices. When coaching support was provided in addition to these techniques, implementation of the novel techniques jumped from sixteen to more than ninety percent (Bush, 1984). In a smaller but similar study, Knight and colleagues reported a large effect size (.96) of coached versus un-coached teachers when compared on a measure of implementation quality (Knight & Cornett, 2009).

Recent studies, mostly qualitative in nature, have served to refine the practice and further define the construct of coaching. Ippolito (2010) identified the need for coaches to

balance responsive and directive behaviors, and documented the amount of feedback, observation, modeling and co-teaching necessary to produce desired teacher and student outcomes. Published interviews, observations and questionnaires have provided information on the components of coaching teachers find most valuable – collaborative support and diversity of strategies - and its impact on the education of both teachers and students (Vanderburg & Stephens, 2010). Marsh, McCombs and Martorell (2010) found that coaches are effective at increasing the use of data-driven instruction and that the coaches' level of effectiveness was tied to the amount of support provided by the administration.

Coaching and Teacher Change

Coaching has been shown to improve teachers' knowledge and application of reading instruction and practices proven to increase reading ability (Armstrong, Cusuamano, Todd & Cohen, 2008; Garet, 2008). Domitrovich, et al. (2009) found that, when provided with regular coaching support and research-based curriculum, Head Start teachers spoke to their students more often and used more complex language in their interactions than comparison teachers. Walpole, McKenna, Uribe-Zarrain and Lamitina (2010) found coaching collaboration, differentiation, and leadership support were significant predictors of teaching improvement in effective reading instruction. Other changes in teacher practices, such as providing increased opportunities for students to respond, increased differentiation of instruction and changes in the classroom literacy

environment have each been attributed to coaching support through scientific study (Carlisle, Cortina & Katz, 2011; Neuman & Wright, 2010; Sailors & Price, 2010).

Coaching and Student Achievement

The lack of scientifically designed studies examining the impact of coaching on student achievement represents the most eminent need in coaching research (Marsh, et al., 2010). Wayne, et al. (2008) suggest the dearth of high quality research in this area is most likely due to the indirect nature of effects of coaching on student outcomes, making correlating student gains with coaching practices difficult. For this reason, caution should be used in examining research citing causal connections. Coaching of educators can only impact learning outcomes of students to the extent it influences educator understanding and application of instructional strategies.

Much of the coaching literature focuses on correlating the amount of time teachers spent receiving coaching support and the related reading performance of students. Shidler (2008) found a significant tie between the amount of time a teacher spent receiving coaching support and the letter identification scores of pre-kindergarten students. Teachers, who spend more hours with coaches than their peers, have yielded students with significantly higher achievement scores and fewer students at risk for reading failure (Ross, 1992; Bean, Hall, Vandermolen & Zigmond, 2010).

Few, if any, studies reported specific coaching components, teacher improvement targets, or reading intervention programs adopted to improve student outcomes. None of

the cited studies applied a measure of coaching or intervention fidelity, making them difficult to replicate and increasingly likely to have results attributed to intervening variables. Thus, further research is necessary to identify the specific components of coaching required to produce significant gains in instructional improvement and student achievement (Knight & Cornett, 2009).

Coaching and Intervention Fidelity

Fidelity—the implementation of a curriculum or intervention in the manner intended—is more closely tied to gains in student achievement than is dosage or frequency of implementation (Kretlow & Bartholomew, 2010). In a review of thirteen coaching studies, Kretlow and Bartholomew (2010) cited significant gains in skill transfer from training to the classroom, when coaching practices were applied. Their study provided “strong evidence for the effectiveness of coaching in promoting the fidelity of evidence-based practices” (Kretlow & Bartholomew, 2010, p. 292). Teemant, Wink and Tyra (2011) indicated that fidelity levels vary with the unique learning and coaching needs of each teacher, but also reported significant gains in skill generalization by participants overall.

Measuring Coaching Effectiveness

Accurately quantifying a moderator variable as dynamic as educator influence is a challenge. Aiming to capture an adequate picture of educator improvement, researchers analyze gains on measures of teacher knowledge, observed frequency or accuracy of

improved instructional practices, as well as the increased performance, engagement or achievement scores of students. As a coach's aim is to empower educators to select and implement interventions proven to improve student achievement, adequately measuring the impact of coaching requires the measurement of teacher change in knowledge, practice, and student outcomes.

Statement of Purpose

The purpose of this study was to examine the extent to which coaching support impacts teacher knowledge, practice and the reading fluency and comprehension abilities of at-risk first grade students. Designed to examine differences in typical professional development and extended coaching support, classrooms were randomized to one of three conditions: a) intervention with typical professional development, b) intervention with professional development and ongoing coaching support, or c) a comparison condition. This study was conducted in an urban area where more than ninety percent of students qualified for free or reduced lunch, with a population of racially and linguistically diverse first graders. Unlike preceding studies, whose findings were limited to correlations between the amount of coaching support and predictive student ability, this study examined the additive effects of coaching support provided to teachers implementing an evidence-based Tier II reading intervention. Multiple measures were used to examine three potential areas of impact: a) teacher knowledge; b) teacher practice as a function of implementation fidelity; and c) student outcomes in reading fluency and comprehension. The coaching model incorporated research-based components of instructional, cognitive

and literacy coaching methods. The following research questions guided the study.

Research Questions

1. When implementing a reading intervention curriculum, did teachers who received professional development training and continued coaching support have significantly better outcomes on a measure of implementation fidelity than teachers who received only professional development training?
2. Did students whose teachers received professional development plus coaching have significantly better outcomes in word reading, reading fluency, and reading comprehension than those of teachers who received professional development only or those who implemented typical school instruction? Were the effects of added coaching support substantively important?
3. Did teachers who received professional development training and ongoing coaching support while implementing a reading intervention curriculum have significantly better outcomes on a survey of teacher knowledge than teachers who only received professional development training?

CHAPTER 2: Review of Literature

Evidence Based Reading Instruction

In 1998, Torgesen stated “the most critical elements of an effective program for the prevention of reading disability at the elementary school level are: a) the right kind and quality of instruction delivered with the b) right level of intensity and duration to c) the right children at the right time” (p. 34). Since that time, the literature on evidence-based reading instruction for at-risk readers has multiplied exponentially. Denton (2012) summarized several instructional components that have been associated with improved reading outcomes for struggling readers. She cited multiple studies which converge on several key ideas; struggling readers achieve the greatest gains when reading instruction: a) tailors explicit instruction with specific and constructive feedback to students; b) provides increased exposure to and practice with connected text both on guided and independent reading levels and c) is engaging and interactive, allowing students multiple opportunities to respond (Denton, 2012).

In 1998, Snow, Burns and Griffin published the National Research Council’s *Preventing Reading Difficulties in Young Children*. Their report highlighted the lack of sufficient progress that students nationwide made in closing the achievement gap, and indicated that approximately forty percent of fourth grade students were not reading at grade level, a statistic that had altered little over the prior twenty-year period. To combat this stagnant growth, the council outlined specific steps in identifying at-risk readers, key predictors of later reading growth, as well as a host of instructional strategies for

intervention in early grades. Among these recommendations were exposing children to both expository and narrative text at a young age, using predictable text with preschoolers, instruction in the alphabetic principle and use of basal readers.

The report of the National Reading Panel (2000) supported many of the same findings and further investigated the most effective content of reading instruction. Their review of hundreds of studies identified phonemic awareness instruction and systematic phonics instruction as the most critical components of an early reading intervention curriculum. Recommendations in the areas of fluency, vocabulary and comprehension were less concise, due to a lack of cohesiveness across experimental studies and the myriad of variables identified across the research corpus. Of note was the correlation between vocabulary acquisition and reading ability as well as the importance of comprehension instruction in early grades in concert with decoding and text reading instruction rather than subsequent to it.

The National Reading Panel and National Research Council reported a scarcity of information on teacher education and training. Although teacher education was a topic selected for inclusion in the report, the Panel cited a lack of scientifically rigorous studies—especially those with both student and teacher outcomes—and produced more questions than conclusions on the topic. Panel members suggested the field is lacking an empirical base to determine the best methods for preparing effective reading instructors. Although professional development was associated with significant gains in student achievement across the few studies reviewed, further research was suggested to address

“how teachers (can) optimally be supported over the long term to ensure implementation of new methods and to ensure student achievement” (National Reading Panel, 2000, p. 17).

Teacher Knowledge and Preparation

With appropriate and early intervention, Torgesen (2000) suggests that half of our hardest-to-teach readers can acquire on-level reading ability, but such gains require significant change in both teacher preparation and practice. Making sufficient reading gains requires placement of our most at-risk readers with well-prepared teachers. Even specially trained teachers are often not adequately prepared to correctly select, deliver and monitor progress of appropriate interventions (Moats, 2009). Students in classrooms with teachers ranked in the top third of performance made gains up to six times greater than students in classes of the lowest performing teachers (Haycock, 2001). Rather than placing the most effective teachers with the students that need them most, the opposite is most often the case. At-risk students are twice as often found in classrooms with under-qualified, inexperienced or uncertified educators than their typical peers (Haycock, 2006). These teachers often lack the necessary knowledge base to serve struggling readers (Lyon & Weiser, 2009). Moats and Foorman (2003) established a moderate correlation between teacher knowledge of reading, overall teaching competence and student achievement, but less than half of the teachers in the four-year longitudinal study were found to have more than a formative understanding of the development and assessment of reading and language. Bos, Mather, Dickson, Podhajski and Chard (2001)

similarly studied in-service and pre-service teacher knowledge of reading instruction and highlighted “a continuing mismatch between what educators believe and know and what convergent research supports as effective early reading instruction for children at risk for reading difficulties” (p. 98). Thus, closing the achievement gap and improving the growth trajectories of at-risk students must begin with improving the knowledge and practice of their in-service educators.

Staff Development

In January of 2002, President Bush enacted new national legislation emphasizing the use of standards-based evaluations and tied federal funding to increased scores on assessments and accountability measures (No Child Left Behind Act of 2001). This comprehensive education reform created an added emphasis on the use of evidence-based interventions and curriculum in the classroom and required teachers have access to professional development. Underperforming districts recruited researchers from major universities, consultants and experienced educators to provide teachers of struggling students with new strategies and skills to improve student outcomes. These investments often resulted in one-day workshops, after-school booster sessions and various other forms of staff development. While useful in bringing new strategies to the teachers, these single-day workshops and professional development sessions have been shown to have little impact in classrooms over time, with the frequency and quality of implementation decreasing rapidly over the course of the school year (Knight, 2007). Further, without follow-up, feedback, and clarification, implementation of newly acquired skills rarely

transfer to the classroom and, therefore, sustainability suffers (Knight, 2009). Such findings gave rise to the role of the coach as a tool to strengthen and sustain the targeted practices learned through typical professional development workshops and take them into the classroom.

Existing Research on Coaching

As a practice, coaching itself is not new to the field of education. Studies indicating the positive effects of peer coaching have been conducted for more than three decades (Showers, Joyce & Bennett, 1987). The cognitive coaching movement shifted the emphasis from peer-evaluation to a more self-reflective process and discussion between professionals with the goal of teacher autonomy (Garmston, Linder & Whitaker, 1993). Most recently, instructional coaching has proven effective in combining the practices of observation, feedback and support to strengthen educator practice (Knight, 2007). Knight and Cornett (n.d.) described the role of instructional coaches as “onsite professional developers who work collaboratively with teachers, empowering them to incorporate research-based methods into their classrooms” (p. 2).

In his study of more than 80 schools across 20 school districts, Bush (1984) found that adding multiple layers of modeling and opportunities produced very little impact compared to that of adding extended coaching support. The International Reading Association standards cite similar research, saying that “professional development, delivered as sustained, job-embedded coaching, maximizes the likelihood that teachers

will translate newly learned skills and strategies into practice” (Neufeld & Roper, 2003; Showers & Joyce, 1996).

Peer Coaching. Peer coaching approaches the task of improving educator instruction with the use of a colleague rather than an outside evaluator. “Peer coaching is the process where teams of teachers regularly observe one another and provide support, companionship, and assistance. Peer coaching programs are non-evaluative, based on the observation of classroom teaching followed by constructive feedback and aimed to improve instruction techniques” (Ackland, 1991, p. 22). Despite the lack of an evaluative component, peer coaching and its sister technique, reciprocal peer coaching, have been shown to have significant positive effects on student outcomes and teacher change (Kohler, Crilley, Shearer & Good, 1997; Zwart, Wubbels, Bergen & Bolhuis, 2007).

Cognitive Coaching. Cognitive coaches take a different approach by focusing on the process behind the product, or the thinking that determines an educator’s choices in the classroom, by teaching them to self-monitor and examine how and why they instruct a certain way. Garmston, Linder and Whitaker (1993, p. 57) describe the process: “cognitive coaching uses a three-phase cycle (pre-conference, observation, and post-conference) to help teachers improve instructional effectiveness by becoming more reflective about teaching.” Edwards and Green (1999) followed thirty-three teachers who implemented cognitive coaching strategies over a three-year period. They found that while teachers’ conversations were more reflective towards the end, the length and language of the conferences changed very little over the course of the study. A self-report

measure demonstrated the social validity of the practice, but not all teachers exhibited growth (Edwards & Green, 1999).

Instructional Coaching. Instructional coaches are “professional developers who work collaboratively with teachers, empowering them to incorporate research-based instructional methods into their classrooms” (Knight & Cornett, n.d., p. 2). Instructional coaching differs from other types of coaching as it involves a collaborative method of examining observation data and provides reciprocity between coaches and teachers, with each learning from the other. In a study of fifty teachers who were taught a new instructional strategy followed by instructional coaching, Cornett and Knight (2009) reported instructional coaching had a large effect size (.96) on the quality with which the teachers implemented the new teaching practices. Their findings also demonstrated that the use of instructional coaching following an initial professional development greatly increased the likelihood that teachers would maintain use of a new strategy, as 15 of 22 teachers reported continued use compared to 3 of the 17 teachers who only attended a workshop (Cornett & Knight, 2009). As another example of its potential, instructional coaching practices produced statistical significance in teacher pedagogy, teacher growth, and classroom organization in a study by Teemant, Wink, and Tyra (2011), who provided 21 elementary school teachers with seven instructional coaching sessions following a multi-day workshop.

Literacy Coaching. Literacy coaches are becoming prevalent in schools as a means to increase student achievement. In year one of a three-year study in Head Start

programs housed in elementary schools, Shidler (2009) found that classrooms that received higher amounts of literacy coaching corresponded with significant student gains in alphabet recognition. Lockwood, McCombs, and Marsh (2010) examined state reading assessment scores of students from 1,000 middle schools to identify a measurable link between a literacy coach and student achievement. Results showed that schools with a state-funded reading coach demonstrated improved average annual reading achievement gains across multiple grades for two of the four cohort years (Lockwood et al., 2010). Matsumara, Sartoris, Bickel, and Garnier (2009) reported a statistically significant positive correlation between the support of school administration and the positive outcomes associated with literacy coaching, which highlights additional factors to consider that may contribute to the effectiveness of a literacy coach.

Coaching Effects on Teacher Knowledge

Carlisle, Cortina and Katz (2011) allowed the teachers themselves to rank improvement in student and teacher learning. In their study, one treatment group received professional development, another received the same training as well as coaching support, while a third group served as a control. Surveys conducted indicated the teachers' perceived changes in their content knowledge more than assessments indicated. They reported that "while perceived knowledge growth between the three groups varied by survey item—no one group notably outperformed another" (Carlisle, Cortina & Katz, 2011). The authors suggested "providing a system for teachers to evaluate their students' learning and their own teaching led to significant gains in teachers' perceptions of

effectiveness as teachers of reading but did not lead to the kind of substantive changes observed in the instructional practices of the KEC (seminar + evaluation + coaching) teachers” (Carlisle, Cortina, & Katz, p. 231). Thus, knowledge and perceived effectiveness did not distinguish groups but use of instructional time on different components (fluency, writing, comprehension) did by the end of the study.

Neuman and Cunningham (2009) reported findings similar to those of Carlisle, Cortina and Katz (2011). Their measure of teacher knowledge revealed that “neither treatment group significantly out-performed the control group on posttest knowledge scores” (p. 553, Neuman & Cunningham, 2009). A third study, by Neuman and Wright (2010) echoed results of previous studies – that coaching support did not demonstrate increases in teacher knowledge of literacy instruction that was required in order to distinguish between the treatment, comparison and control groups.

In an effort to answer the question of whether or not coaching support increases teacher knowledge, all three studies used self-report or researcher-made measurements to assess growth in teacher understanding of literacy and reading principles. None of the studies presented here indicated significant growth in content knowledge of coached teachers over their non-coached peers. In fact, researchers in all three studies reported scores on knowledge assessments could not distinguish coached or un-coached teachers from control groups who did not receive professional development. This finding calls into question our aim in providing in-service teacher education and requires us to examine whether our goal is to improve teacher knowledge or practice. As the answer

may not be mutually exclusive, the fact that professional development practices are divided between these two goals is worth further exploration. Current research indicates our initial efforts as teacher educators may be best placed in the improvement of instructional practice, rather than the simple dissemination of information usually provided in one-time workshops and trainings.

Coaching Effects on Teaching Practices

Carlisle, Cortina, and Katz (2011) studied the impact of increasing levels of professional development support on cohorts of first grade teachers in schools eligible for or receiving Reading First funding. In their additive model, one-third of their sample (n=35) attended nine professional development seminars to learn new reading and spelling instruction strategies. A second group attended the same seminars but were additionally taught and provided minimal interventionist support to evaluate the effectiveness of their implementation, using student scores, and to adapt instruction accordingly. The third cohort (n=43) added to the previous components' regularly planned collaboration and consultation with a literacy coach. This additive condition (seminar + evaluations + literacy coaching) was found to be significant in creating teacher change, as teachers in this condition were found to use the comprehension program during more of the instructional period and to provide a wider variety of literacy activities to their students.

Using a single-subject design to more closely examine individual teacher change, Hsieh, Hemmeter, McCollum and Ostrosky (2009) grouped emerging literacy skills to be taught into three clusters: (a) oral language and comprehension, (b) phonological awareness and alphabetic principle, (c) print concepts and written language. They provided separate coaching support for each component. After six weeks of intervention, five observed teachers demonstrated an increased use of emergent literacy teaching strategies across all participants with markedly higher use of literacy strategies during coaching periods. Teachers already using some literacy strategies in their teaching repertoire did so with more consistency and variety with coaching support than during baseline. Maintenance scores indicated teachers continued to use an increased numbers of strategies even weeks after the coaching support was discontinued.

In another study, Sailors and Price (2010) provided 27 teachers a two-day summer inservice training introducing the use of cognitive reading strategies in the classroom. Their peers (n=17) received the same workshop and additional support from one of two trained reading coaches who provided modeling, co-teaching and feedback focused on the cognitive reading and decoding strategies. Researchers used the Comprehension Instruction Observation Protocol System to measure the extent to which teachers implemented the information introduced during the professional development. Observation data revealed that coached teachers provided a statistically significant increase in opportunities to use reading strategies and engage in comprehension activities ($d=.78$). Coached teachers also used more structured explanations of targeted strategies

($d=.64$). The authors concluded the coaching model used in their study was equally effective with teachers despite differences in levels of education, teaching experience and across multiple subjects and grade levels.

To examine the effects of coaching on the pre-literacy instruction of early childhood educators, Neuman and Cunningham (2009) used two treatment groups and a control group in several large urban areas with students considered to be “high priority.” All treatment teachers participated in a course on improving language and literacy instruction with students three to five years old. One half of these teachers ($n= 85$) received weekly support from one of fourteen assigned coaches who were in their classrooms for an hour or more per week. Observation instruments (ELLCO or CHELLO) showed that teachers receiving only the professional development training had “negligible effects” in regards to their instruction, but teachers who received the added benefit of coaching support had significantly higher posttest classroom observation scores. Subtest scores indicated teacher support for learning and several environmental arrangement indicators allowed students more access to literacy materials and opportunities for interaction with writing and print materials.

A similar study by Neuman and Wright (2010) examined schools serving pre-school students from low-income backgrounds whose teachers participated in thirty hours of coursework, after which half of the teachers received weekly on-site coaching support for three hours per visit. The authors identified statistically significant improvements to the literacy environment in classrooms of coached teachers, changes that were still in

effect five months after coaching support ended. The authors suggest these improvements support the use of coaching over traditional professional development when compared with teachers who only received information through a training course. Those teachers did not exhibit distinguishable differences in ratings of classroom environment or instructional practices.

Matsumara (2010) found that despite the fact that teachers engaged with their coaches less frequently than researchers intended, self-report surveys indicated greater participation in literacy coaching activities and statistically significant improvements in four areas by coach-supported teachers over their non-coached peers. Differentiation of instruction, planning and reflecting on instruction, increased knowledge of theory behind reading comprehension instruction, and support during lessons present evidence of positive and statistically significant change in teacher behaviors when provided coaching support. Multiple studies suggest coached teachers are more likely to differentiate instruction (Carlisle, Cortina, & Katz, 2011) and improve access to and use of literacy materials (Carlisle, Cortina & Katz, 2011, Neueman & Wright, 2010). Coached teachers were also found to use a wider variety of literacy teaching strategies and instruction (Hsieh, Hemmeter, McCollum & Ostrosky, 2009; Sailors & Price, 2010) as well as increased opportunities to engage in comprehension activities (Sailors & Price, 2010). This supports the research of Cornett and Knight (2009), Joyce and Showers (2002), Kohler (1997), and Bush (1984) that asserts that coached teachers implement changes in instructional practice more readily than their un-coached peers.

Coaching Effects on Student Reading Outcomes

Sailors and Price (2010) reported that students of teachers participating with a coach had higher reading achievement scores and were more engaged in reading activities than students of non-coached teachers. An examination of pre and posttest scores on the Group Reading Assessment and Diagnostic Evaluation (GRADE) assessment indicated small but significant ($d=.33$) effects for student reading achievement scores within the intervention group. Tying this result to the coached teachers, researchers reported “71% of the students whose teachers displayed characteristics specific to the composition of the comprehension strategies variable resulted in positive GRADE score changes” (Sailors & Price, 2010, p. 315). The authors suggested that the increase in scores was most likely due to coached teachers providing more opportunities to respond and engage in comprehension activities than their non-coached peers. Important to the interpretation is the limited amount of time the teachers were coached across the school year, an average of 329 minutes (less than six hours). Even so, the study yielded small but significant results, suggesting that even a small investment in classroom coaching can produce significant changes in both teacher practice and student reading achievement.

To examine outcomes on a school-wide level, Matsumara (2010) and colleagues collected baseline data on teaching instruction and student performance prior to training teachers in treatment and comparison schools on a “Question the Author” reading comprehension strategy. Schools whose teachers actively engaged in coaching support

had significantly higher school-level gains on the state standardized assessment for English Language Learners (ELLs). Whereas there were not significant effects across all types of students on the state assessment, the English Language Learners averaged almost 60 points higher than those in schools where teachers did not have coaching support.

Using students' year-one scores on DIBELS and Terra Nova Reading assessments as baseline data, Biancarosa, Bryk and Dexter (2010) conducted a four-year longitudinal study of teachers and students in seventeen schools where at least one teacher was trained and then employed as a literacy coach. Results were significant for "increasing improvements" in student literacy growth (over their baseline growth rates in year one) across the three years of implementation with effect sizes of .22, .37, and .47 for years two, three and four respectively. Authors reported that students who entered the study with lower rates of literacy increased abilities at a faster rate than their initially higher-achieving peers and that this growth was maintained over summer breaks in instruction. By year four, these researchers report students in school where literacy coaching was provided to teachers were learning 32 percent more than the year prior to implementation.

Cusumano, et al. (2006) applied coaching support to one half of an intervention group of early childhood teachers, all of whom received a fifteen-week college course addressing the teaching of literacy skills to young children. Whereas there was significant increase in alliteration scores found in students of teachers who took the course, it could not be directly contributed to coaching support. As results were contrary to their hypothesis, the authors suggested the shortened length of coaching support provided,

only four to five weeks, may not have been sufficient to demonstrate significance and suggested further study of the aspects of coaching required to bring about effective change in teacher behavior.

Each of these four studies demonstrated that change in teacher practice can correspond to increased student scores on both distal and proximal measures of literacy and reading skills (Biancarosa, Bryk, & Dexter, 2010; Cusumano, Armstrong, Cohen & Todd, 2006; Matsumara, Garnier, Correnti, Junker & Bickel, 2010; Sailors & Price; 2010). Students made measurable and statistically significant gains in alliteration, phonemic awareness tasks, reading and comprehension measures on progress monitoring, grade-level and state-wide assessments. These findings support those of others (Cornett & Knight, 2009; Neufield & Roper, 2003; Teemant, Wink & Tyra, 2011), that coaching can lead to improved teacher practice and significant student gains in reading achievement.

Summary of Coaching Research

Findings from these studies supported the use of coaching as an effective way to improve teacher practice, indicating this type of sustained professional development and instructional support produces statistically significant increases in teachers' use of instructional time, differentiated instruction, and increased opportunities to engage in literacy and comprehension activities. Although improvements in teacher knowledge and theory were not statistically significant, the studies clearly demonstrated improved

teacher practice when coaching support was provided. Coached teachers not only used a greater number of literacy strategies throughout their instruction, they also used a greater variety of strategies and did so more consistently. These improved instructional practices were maintained after coaching support was removed. Students of coached teachers were found to be more engaged in reading activities, have higher reading achievement scores on state and standardized measures and demonstrate improved rates of growth on reading-related measures.

Coaching For Fidelity of Evidenced Based Practices

In 2010, Kretlow and Bartholomew conducted a comprehensive review of research to identify the impact of coaching on changes in pre-service and in-service teachers' implementation of evidence-based practices. This synthesis of 13 quantitative studies concluded that multiple evidence-based teaching practices improved when teachers were provided coaching, including increased labeled praise. In these cases, direct instruction practices occurred more frequently and teachers showed increased adaptation of academic materials. Coached teachers included in the synthesis (n = 110) were also more likely to prompt students for target behavior and found that student engagement and responses increased in these classrooms while inappropriate behavior decreased.

Authors of the review concluded that coaching interventions led to improvements in teaching accuracy, although accuracy was defined differently across the studies (Kretlow & Bartholomew, 2010). In-depth interviews were conducted in eight studies to

determine social validity of coaching practices. Teachers in all eight studies reported positive experiences. While these findings are encouraging, researchers admit, “Only five studies in this review reported any student data and of those five, only three demonstrated experimental effects of coaching and improvements in student outcomes” (Kretlow & Bartholomew, 2010, p. 293).

Summary

With the understanding that teacher knowledge, certification and preparation are highly correlated to student achievement (Darling-Hammond, 1999) and the added indication that stand-alone professional development workshops result in ten percent or lower implementation rates (Knight, 2009), it behooves us as educators to increase the likelihood of student success by bringing evidence-based practices into classrooms and providing the sustained support required for effective implementation. Snow, Griffin and Burns (2005) suggest that to create true change we must go beyond simply trying to expand educators’ content knowledge and impact their pedagogical knowledge. In essence, we cannot expect lasting change in instruction, change that will impact student understanding and acquisition of skills, by merely teaching teachers about reading. We must demonstrate, model, work side-by-side, plan, evaluate, reflect and identify targets for pedagogical improvement. Lasting change and sustained instructional improvement involves bringing the correct “what” into the classroom in the form of evidence-based practices, but also the correct “how” and “why” of effective instruction.

In reference to the lengthy amount of time it takes to bring research-proven techniques into contact with students most in need of them, Foorman and Moats (2004) reported, “We know how to implement sound research-based practices in early reading. We are learning what needs to be in place to sustain them” (p.54). Their assertion echoes the sentiments of others who have cited the need for scientifically structured studies that simultaneously examine teacher and student effects of sustained professional development – and the coaching required to maintain implementation of novel practices with fidelity.

CHAPTER THREE: Method

Overview

At-risk readers require increased opportunities to engage with expository and narrative text, explicit and systematic instruction in phonics and phonemic awareness and direct teaching of comprehension skills and strategies (Foorman & Moats, 2004; Foorman & Torgesen, 2001). Implemented accurately, evidence-based intervention strategies have been shown to improve the reading ability of both at-risk readers and those with identified reading disabilities (Torgesen, et al, 2001). Mastery of phonemic awareness, decoding and reading comprehension skills are highly predictive of later school success. When students leave early elementary grades with these skills unmastered, they are presented with fewer opportunities to engage with instructionally appropriate text, leaving them less likely to acquire grade-level reading abilities over time (Torgesen, 2004).

Until recently, evidence-based intervention strategies were not commonly implemented in general education classes, as they were most-often applied in pull-out resource or other support settings. The response-to-intervention framework, designed to provide intervention support earlier in the career of struggling students, gives general education teachers the responsibility of instructing our most at-risk readers. Many of these teachers are ill prepared to provide the type of instruction necessary to close the achievement gap between these students and their typically developing peers (Haycock, 2011).

Efforts to improve instructional practice and teacher content knowledge—and thus prepare them with effective interventions for struggling students—are most often attempted through professional development workshops and trainings. Unfortunately, research indicates that without sustained modeling, feedback and observation these practices are implemented as infrequently as ten percent of the time (Knight, 2009). When coaching support is provided to educators, use of evidence-based practices with fidelity increases and educators are more apt to use data to influence their instructional decisions and differentiate instruction (Carlisle, Cortina & Katz, 2011; Lockwood, McCombs & Marsh, 2010; Kretlow & Bartholomew, 2010).

Research Questions

The purpose of this study was to examine the extent to which coaching support provided to first grade teachers using an evidence-based reading intervention program impacted teacher knowledge, practice and student reading outcomes. The study was conducted as part of a larger pilot study aimed at establishing preliminary evidence for the reading intervention curriculum. The purpose of an exploratory study is to examine data to determine potential hypothesis that should be further explored through more rigorous, replicated studies (Schochet, 2008). As part of an exploratory study, this study was conducted with a relatively small number of participants, resulting in insufficient statistical power to provide conclusive evidence as to specific teaching or coaching practices. The study addressed the following research questions:

1. When implementing a reading intervention curriculum did teachers who received professional development training and continued coaching support have significantly better outcomes on a measure of implementation fidelity than teachers who received only professional development training?
2. Did students whose teachers received professional development plus coaching have significantly better outcomes in word reading, reading fluency, and reading comprehension than those of teachers who received professional development only or those who implemented typical school instruction? Were the effects of added coaching support substantively important?
3. Did teachers who received professional development training and ongoing coaching support while implementing a reading intervention curriculum have significantly better outcomes on a survey of teacher knowledge than teachers who only received professional development training?

Coaching support was provided to teachers implementing evidence-based reading strategies as part of a Tier 2 intervention curriculum designed to increase the comprehension and decoding abilities of at-risk first grade readers. To determine the extent to which coaching support impacted teacher knowledge, practice and student outcomes of participants, first grade classroom teachers were randomly assigned to one of three groups: (a) a group that received an evidence-based intervention curriculum, professional development (PD) in its implementation, and bi-weekly coaching support (PD plus coaching); (b) a group that received the intervention curriculum and training,

but no coaching (PD only); or (c) a comparison group that did not receive the curriculum, training, or coaching, but implemented typical school instruction (comparison group). All PD plus coaching and PD only teachers received a two-day professional development training to prepare them to deliver the reading curriculum as intended prior to implementation. Two follow-up PD sessions were provided approximately four and eight weeks after the start of the intervention for all teachers (PD only and PD plus coaching) who implemented the reading curriculum. In the PD plus coaching condition, coaches supported teacher implementation of research-supported instructional strategies for at-risk readers using observation, modeling, co-teaching, co-planning, and feedback on collaboratively selected areas targeted for improvement. The PD only group participated in the same professional development on the implementation of the intervention curriculum (two PD days prior to intervention and two PD days during the intervention period). PD-only participants were also able to email or call research personnel for assistance throughout the course of the intervention but were not provided in-class coaching support. The comparison group continued to implement the district-adopted reading curriculum as set forth by their administration.

Both teacher and student outcomes were measured. To examine changes in teacher practice, videotaped lessons were coded for implementation fidelity for teachers in the PD only and PD plus coaching groups. A percent perfect (number of points earned divided by number of points possible) fidelity score of all presented components was computed for each teacher implementing the reading curriculum. Scores of teachers in

the PD only and PD plus coaching groups were then compared and examined for statistically significant differences and effect sizes were calculated. Using the standards set forth by the What Works Clearinghouse, an effect size of .25 was considered substantively important (What Works Clearinghouse, Procedures and Standards Handbook, v2.1). To examine the impact that coaching teachers had on their students' reading ability, students' word reading, reading fluency and reading comprehension scores on a battery of standardized and researcher-created measures were collected and compared across the three conditions. Finally, a teacher knowledge survey was administered to teachers in each of the three conditions and used as an exploratory measure to examine changes in teachers' pedagogical knowledge of evidence-based instruction for at-risk readers prior to and following the intervention.

Research Design

This study was conducted as part of a three-year project funded by a Development and Innovation grant from the Institute of Education Sciences to demonstrate the feasibility, usability and promise of a Tier 2 intervention curriculum designed for first grade students at risk for reading difficulties in both decoding and comprehension. Using an experimental pre-posttest design, teachers in participating schools were randomly assigned to one of three conditions: PD only (n=7), PD plus coaching support (n=6), or typical school instruction (TSI), a comparison condition (n=8).

A treatment/comparison design was selected to minimize threats to internal validity such as maturation, history, attrition and selection bias. Three teachers were

randomized to PD only, PD plus coaching, and comparison conditions at each of six participating schools. Two campuses had only two participating teachers, so teachers at these campuses were paired prior to randomization. One teacher at each of these schools was randomized to the comparison condition with the remaining two teachers randomized (one each) to the PD-only and PD plus coaching conditions. This process resulted in seven teachers each in the PD-only, six teachers in the PD plus coaching condition and eight teachers in the comparison condition. Because teachers were equally likely to be randomized into each of the three conditions, selection bias was minimized. The use of a control group also helped guard against the threat of maturation as students in both treatment and control conditions could be expected to mature at similar rates. To guard against expectancy effects, data collectors and assessors were blind to participants' assigned conditions.

To optimize the use of resources and research personnel, the eight participating schools were divided into two groups. One group began the reading curriculum intervention one week later than the first group. Pre- and post-test assessments were administered one week later in the second group and coaching support also followed this one-week lag. This design is described in Figure 1.

Figure 1. Study Design.

STUDY WEEK	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
GROUP 1: 10 teachers (3 PD only, 4 PD plus coaching, 3 comparison)	Intervention		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17					
	Assessments																				Post			
	Teachers		TKS PD			FI				FI						PD			FI				TKS	
GROUP 2: 11 teachers (4 PD only, 3 PD plus coaching, 4 comparison)	Intervention			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17				
	Assessments			Pretest																		Post		
	Teacher Knowledge		TKS PD				FI				PD		FI				PD			FI				TKS

Notes: D= DIBELS; T=TOSREC; FG= Focus Group (intervention teachers only); TI=Teacher Interview (all teachers); FI=Fidelity of Implementation Observations; TKS=Teacher Knowledge Survey Denton, C. and Solari, E. (2012).

Participants

Teachers. Twenty-one teachers assigned to first-grade classrooms participated in the study. Demographic information including ethnicity, educational background, years of teaching experience, and number of years teaching students with reading difficulties was obtained for intervention and comparison teachers. All participating teachers were female. Of the twenty-one participants, ten were Caucasian, nine were African American, one was Asian American and one was Hispanic. The highest degree obtained by one teacher was a doctoral degree, seven teachers held Master’s degrees and thirteen held Bachelor’s degrees. Two teachers held additional certifications as reading specialists. The average number of years of teaching experience held by coached and un-coached teachers were 11 and 10, respectively. Teachers in both groups also reported similar number of years teaching students with reading difficulties (coached average = 10.2, un-coached average = 9.1). Teachers randomized to the comparison condition reported markedly fewer years of teacher experience and consequently fewer years teaching students with reading difficulties than those randomized to the intervention. Demographic information is reported in Table 1.

Table 1.
Teacher Demographic Information.

	Ethnicity	Degrees Held	Average Number Years of Teaching	Average Years Teaching Students with Reading Difficulties
Coached (n=6)	4 C, 2 AA	4 Bachelor’s 2 Master’s	11(1-32)	10.2(0-30)
Un-coached (n=7)	3 C, 2 AA, 1 AS, 1 H	5 Bachelor’s 2 Master’s	10(1-17)	9.1 (2-15)
Comparison (n=8)	3 C, 5 AA	4 Bachelor’s 3 Master’s 1 Doctoral	7 (1-20)	6.5 (0-20)

**Note:* C=Caucasian, AA=African American, AS=Asian American, H=Hispanic

Students. Ninety-nine students were pre-tested at the beginning of the study. The sample included fifteen classrooms of five student participants each and six classrooms in which four students participated. Over the course of the intervention, eight students moved or were withdrawn from school, leaving ninety-one students for which there were both pre- and post-test measures. Demographic information including race, gender, exceptionality and qualification for free or reduced lunch was gathered for each participating student and is reported in Table 2.

Table 2.

Student Demographic Information.

	Coached (n=27)	Un-coached (n=29)	Comparison (n=35)
Gender			
Male	48	59	54
Female	52	41	46
Special Education Designation	7	0	11
Free & Reduced Lunch	85	100	100
Limited English Proficiency	41	17	23
Ethnicity			
African American	26	48	37
Hispanic	66	41	51
Asian American	4	0	0
Caucasian	0	10	6
Other	4	0	6

Note: All numbers represent percentages.

Coaches. Two trained research assistants served as coaches. Coaches each had a minimum of eight years prior experience teaching at-risk students, held Master’s degrees in education-related fields and had two or more years of experience coaching educators. Both coaches previously worked as writers of the intervention curriculum, and thus were familiar with the content and proper implementation of the lessons. Coaches were assigned to intervention teachers early in the study and remained with those teachers for the remainder of the intervention.

Setting

The study was conducted in a large Southwestern school district in an urban area. The district served 203,000 students, more than half of which were enrolled in one of 181 elementary schools. As part of highly diverse urban metropolitan area, district attendees were 61.7 percent Hispanic, 26.5 percent African-American, 7.8 percent Caucasian, 2.9 percent Asian and 1.1 percent American Indian, Alaskan Native, Hawaiian/Islander or reported a combination of more than one race. More than ninety-five percent of the students attended schools which qualified for Title I funding and 63.2 percent of the district population had been identified as at-risk in one or more academic areas. Just over one-fifth of the district (20.6%) was bilingual, with closer to one-third (30.7%) designated as Limited English Proficient. Just over eight percent of the district's students received special education services. Academically, the district consistently performed below the state average in all five content areas on statewide assessments.

The eight participating schools were spread across the district geographically and campus demographics closely reflected those of the district as a whole. Six of the eight campuses were public schools and two were charter schools associated with the Knowledge Is Power Program (KIPP). All included schools received Title I funding.

Student assessments were administered in a quiet area on each campus designated by administrative personnel (e.g., library, conference room). All intervention sessions and observations were conducted within treatment teachers' classrooms during regularly scheduled class periods. Coaching sessions were also conducted within the teachers' classrooms or another teacher-selected area. Feedback sessions occurred during a time of the teachers' choosing.

Participant Selection

Teachers. Research personnel obtained district-level approval for study implementation prior to contacting the administrative staff of individual schools. Principals were then contacted by intervention personnel who described findings from prior pilot studies, time requirements and obligations for potential participants. Administrators who agreed to have their schools included in the study were then asked to invite first grade teachers from their campus to attend an introductory meeting conducted by research personnel wherein the content of the first-grade tier-two intervention curriculum was described.

To participate, teachers and students had to be assigned to typical first-grade classrooms. Teachers of bilingual, self-contained special education, deaf and hard-of-hearing classes or classes designated for students with severe emotional, behavioral or developmental disabilities were not invited to participate in the study. Informed consent was obtained for all participating educators.

Students. Participating teachers identified the eight children performing most poorly in reading within their respective classrooms. Classroom teachers then administered the district-mandated Texas Primary Reading Inventory (TPRI) and a researcher-created measure of word-reading accuracy. Students' word-reading scores were then multiplied by two and added to their TPRI comprehension score to create a composite score. The five students with the lowest composite score in each class qualified for participation in the study and were then discussed with the teacher for final approval. Research personnel met with the teachers in each class to present the students who qualified for the study, teachers then had the ability to agree with the selection or present a case for the substitution of another child. After the teacher and research personnel agreed upon the target students, students were given permission forms to take home for parents to

sign, indicating consent for their child to be involved in the study. All children also signed assent forms prior to pre-testing.

Procedure and Conditions

Procedure. The study included the following steps: a) students were screened; b) teachers were trained by project personnel; c) students and teachers were pre-tested; d) intervention curriculum was implemented; e) coaching support was provided to identified teachers; f) fidelity observations of teachers implementing the reading curriculum were conducted; and g) students and teachers were post-tested.

PD only condition. Intervention teachers in the PD-only condition participated in the initial two-day professional development training prior to implementing the reading curriculum. Teachers assigned to this condition also participated in two follow-up professional development sessions, one of which included teachers spending a day observing two teachers who were implementing the curriculum and discussing their observations with research staff. These teachers were able to ask questions and received support via email or phone from research personnel over the course of the intervention, but did not receive in-class visits.

PD plus coaching. Teachers assigned to the PD + coaching condition also received the initial two-day professional development training as well as the two follow-up professional development trainings, including one on-site observation of the two teachers implementing the curriculum. In addition, these participants received two days of coaching support from one of two trained coaches every other week.

Typical school instruction (TPI) condition. Teachers randomized to the typical school instruction (TSI) or comparison condition continued to instruct their students using the district-selected curriculum. These teachers also identified eight students with poor reading ability and used the same screening instruments as were used with the treatment students to select the five lowest performing students in their respective classes. The five selected students received daily reading instruction according to typical practice in their schools as well as any additional tutoring or supplemental school instruction for which they qualified. Comparison teachers did not participate in the professional development trainings or receive any instructional support from research personnel. Pre- and post-test assessments were administered for comparison purposes.

Interventions

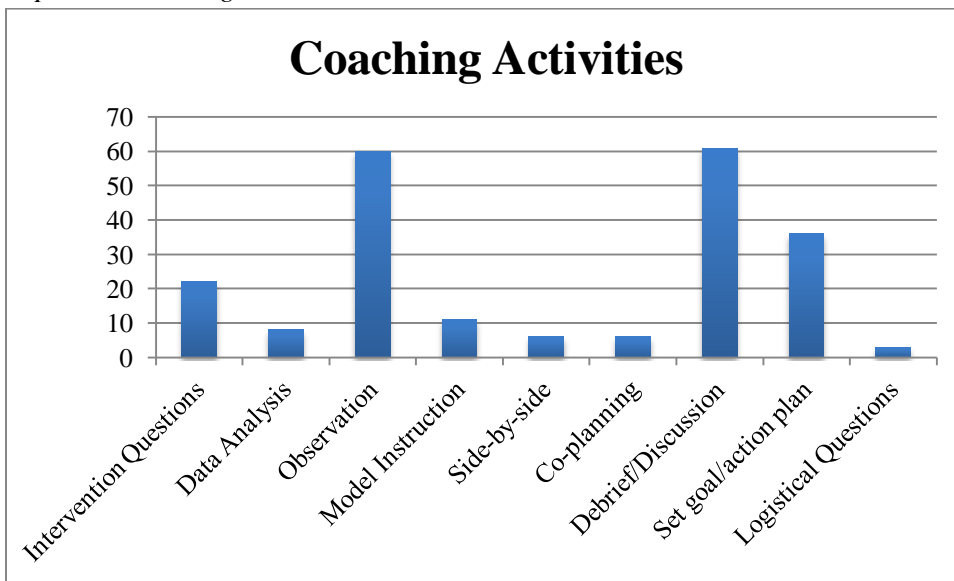
Coaching Intervention. Coaching support was provided to the six teachers randomized to this condition. Coaches scheduled initial observations during the second week of the implementation of the new curriculum. The reading intervention included a fifteen minute whole class comprehension lesson followed by a thirty minute small group lesson on comprehension, word-attack and text-reading skills. During each visit, the coaches observed both a whole group and small group lesson and met with teachers to discuss progress and any problems they were having implementing the curriculum. Teachers and coaches co-selected target areas for improvement and identified the type of support that would prove most beneficial for the following visit (modeling, side-by-side teaching, additional observation).

During the remaining visits, the coaches monitored levels of proficiency of specific instructional practices including: a) the extent to which the teacher followed the curriculum as written; b) the effective use of instructional activities; c) appropriate scaffolding and feedback provided to students; d) methods used to keep students engaged and on-task; e) pacing and use of instructional time; and f) evidence of teacher preparation.

Following each observation, the coach and teacher met together during a planning period (or other teacher-selected time) to discuss lesson execution. During each visit, the coach asked the teacher to identify strengths and weaknesses of instruction. Using reflective listening and questioning strategies drawn from instructional, cognitive and literacy coaching research, the coaches guided conversations, modeled instruction, and provided scaffolding to promote deeper thinking, support problem solving strategies and improve teachers' data analysis skills.

Coaches visited teachers once every two weeks on average as teachers' schedules allowed. Teachers received an average of ten coaching sessions over the seventeen-week intervention period (range 7-12). Observations were usually 45 minutes in length with a 30-minute follow-up debriefing/planning session. Coaches reported the following activities and their frequency: discussion of questions regarding the implementation of the intervention (22), discussion of data analysis (8), observation (60), modeling of instruction (11), side-by-side teaching (6), co-planning (6), debriefing (61), setting goals (36), and/or answering logistical questions (3). Figure 2 illustrates the number of times coaches reported completing an activity over the course of the intervention.

Figure 2.
Reported Coaching Activities.



Reading Intervention. The Reading RULES! curriculum was developed by teachers and researchers as an intervention for at-risk readers needing Tier 2 support. This study was conducted during the final year of a three-year Development and Innovation project supported by funds from the Institute of Educational Sciences. Field tests were conducted during two prior years during which qualitative (feedback from teachers) and quantitative data (student scores) were used to help determine elements of the curriculum deemed most effective and those needing revision.

The Reading RULES curriculum provided explicit, systematic instruction in word study, text reading and comprehension. The intervention was designed to be implemented four times per week for a total of forty-five minutes per day. Each forty-five minute lesson was divided into a fifteen minute whole-group comprehension component and a thirty-minute small-group intervention component that included additional practice in comprehension, word study and text

reading. Although the curriculum was scripted, teachers were encouraged to learn and apply the routines taught in each activity using the script as a guide rather than reading it verbatim.

- a) Whole group comprehension (10-15 minutes) – Conducted with the entire class, whole group comprehension lessons used teacher-read texts and scripted lessons to teach six skills in a sequential order: 1) direct recall; 2) background knowledge and making connections; 3) narrative story elements and retelling; 4) asking questions; 5) making inferences; and 6) main idea. Each skill was taught explicitly during an initial anchor lesson where the teacher used modeling, guided and independent student practice to demonstrate and allow students to practice the new skill. Following lessons in the same unit used the same or different text in a read-aloud format. Read-aloud lessons followed a before, during and after reading routine where teachers introduced the text, reminded students of the comprehension skill they were practicing and used a guiding question to focus students on the targeted skill. During reading, the teacher stopped at pre-selected points in the text to introduce new vocabulary and ask scripted comprehension questions. Midway through the story, the teacher also reminded the students to listen for the answer to the guiding question. After reading, the teacher led a discussion about the guiding question and reviewed the comprehension skill once more. For each comprehension skill, students were taught a hand signal to help them recall the skill they were practicing.
- b) Small group comprehension (10 minutes) – Building on the whole-group comprehension lesson, the small group lessons were designed and implemented to reteach and provide additional opportunities for practice of the newly learned skill. Manipulatives and pictures from the book were provided for students to practice the six comprehension skills in the curriculum (see above).

- c) Word study (10 minutes) – Each day a group of five to six activities were used to build reading skills in the following areas:
- a. Phonemic Awareness – students used a variety of activities and materials to complete auditory discrimination and elision tasks.
 - b. Sound Analysis – students learned to map sounds to print through activities in which they substituted sounds, stretched and spelled words using previously taught letter-sounds.
 - c. Decoding- students were provided opportunities to practice reading simple CVC words, words with consonant blends, words with special endings, and multisyllabic words.
 - d. Letter Sound Instruction – students were taught to recognize consonants, vowels, digraphs, additional letter combinations and their associated sounds.
 - e. High Frequency Words- students were introduced to new high frequency words and review previously learned words with a focus on identifying words accurately and fluently.
- d) Text Reading (10 minutes) – Text reading lessons combined skills students learned in previous comprehension and word study lessons and applied them to decodable text. During early units, students read simple sentences before moving to small illustrated books with decodable text. Students learned to identify unknown words in text by applying a three-step strategy: (a) look for parts you know; (b) sound it out; and (c) check it!.

Before reading, the teacher posed a guiding question to students. Each day individual students were selected to read several pages aloud while the teacher provided

necessary feedback and scaffolding. The teacher rotated the selected student so that each student had ample opportunity to practice reading decodable text. As the child read, the teacher was asked to mark certain places in the text where additional praise or teaching should occur. Each text was designed to provide practice in a specific letter-sound pattern and high-frequency words. Texts were arranged to build on previously taught letter sounds and words. After reading, the teacher was asked to revisit the guiding question and use the marked praise and teaching points to provide additional feedback.

- e) Fluency (2-3 minutes) - After students initially read a text, students engaged in one of five fluency strategies to increase reading fluency: a) read it like talking; b) read it with punctuation; 3) read it with a friend; 4) read it in phrases; or 5) beat the clock.

Professional Development (Intervention) Research personnel provided intervention teachers with a two-day training to demonstrate the instructional components and appropriate implementation of the Reading RULES! curriculum. The two principal investigators and three research personnel modeled whole and small group comprehension lessons, text reading lessons and word study activities and provided opportunities for guided and independent practice. Teachers received instruction on how to administer mastery tests designed to help teachers place students at the appropriate level for word study and text reading activities and on providing scaffolding and feedback support during text reading instruction. Participants were also shown video segments of teachers implementing the curriculum with first grade students. During the two eight-hour sessions, teachers were given structured time to practice lessons with peers, ask questions about content, and become familiar with materials.

Two follow-up professional development sessions were conducted approximately four and eight weeks after the start of the intervention, respectively. These additional two eight-hour

trainings provided researchers time to model and practice activities that occurred only in higher level units. One of these two follow-up professional development trainings allowed teachers the opportunity to visit a campus and observe two teachers who implemented the curriculum well. After these observations, teachers met with research personnel to discuss adaptations they might make to their own methods to increase student engagement and implementation accuracy. Teachers were also able to ask questions and problem solve regarding implementation concerns that they were experiencing at the time.

Training

Coaches. Prior to their participation in the study, both coaches had multiple years of experience and training in the area of educational coaching. Coaches also worked closely with the project coordinator and principal investigator, who each previously worked as educational coaches, to discuss methods and coaching models. Coaches also previously rated taped sessions of prior years' implementation to refine, validate and obtain reliability on the Fidelity of Implementation instrument. Both coaches previously served as writers and professional development trainers of the curriculum and thus were highly familiar with the intervention materials.

Assessors. Prior to administering assessments, assessors were trained by an assessment coordinator assigned to the grant. An assessment binder was compiled with directions for obtaining child assent, administration of each assessment, proper establishment of basals and ceilings as well as scoring procedures. Assessors attended a six-hour training session prior to the start of the study, where administration of measures were modeled and trainees were allowed to practice administration of items with research personnel. Trainees were then given a full day to

practice administration at home prior to returning for a check-out procedure. Assessors who did not administer items accurately were retrained and asked to retest prior to assessing students.

Measures

Coaching Measures

Coaching Fidelity (Appendix A) The Reading Rules Coaching Fidelity Form was created to measure coaches' adherence to the prescribed coaching intervention components. Components were rated as present, partially present, not present, or not applicable. The ten components rated on the measure were: a) coach observed entire lesson; b) coach collected student/teacher data; c) coach provided in-lesson support, scaffolding or modeling; d) coach used data to inform feedback; e) coach provided specific, positive feedback; f) coach indicated areas to be strengthened or improved; g) coach listened to teacher concerns/questions; h) coach addressed questions/concerns and provided resources; i) coach and teacher co-selected a target/goal for improvement; and j) coached suggested action steps to meet goal/target. Each of the two coaches was observed twice during the intervention – once each with two different teachers – by the project coordinator.

Teacher Measures

Teacher Practice. (Appendix B) The Reading Rules Fidelity and Quality of Implementation Instrument was used at three points over the course of the intervention to rate the quality and quantity of curriculum components implemented in each observed lesson. The instrument was developed by research personnel and validated through observations of prior years' participants. Observers indicated the number of minutes spent on each of the required

lesson components and rated each component on a three point Likert-type scale, with one indicating low and three indicating high fidelity. Each instructional component (e.g., whole group comprehension, word study) was rated for fidelity across four categories: a) oral presentation of script; b) implementation of instructional activities; c) scaffolding and feedback; and d) accuracy of student responses.

Teacher fidelity ratings were calculated by dividing the number of points earned by the total number of points possible. Fidelity scores were computed for each teacher, and scores of teachers in each condition (PD only, PD plus coaching and comparison conditions) were then averaged and compared, looking for statistically significant differences and effect sizes to determine the impact of coaching support on teacher practices. Results are reported in chapter four.

Teacher Knowledge. (Appendix C) A researcher-made Teacher Knowledge Survey was administered to participating teachers prior to the initial professional development training and the week following the intervention. Designed to measure teachers' understanding of reading instruction for at-risk students, the assessment consisted of 22 items related to evidence-based intervention practices. The first ten items were statements including common misconceptions regarding effective reading instruction. Teachers were asked to rate each statement using a Likert-type scale and indicate whether they 1) strongly disagree; 2) disagree; 3) neither agree or disagree; 4) agree; or 5) strongly disagree. The remaining twelve items were open-ended questions, case scenarios and multiple choice questions; the responses to these questions were designed to serve as indicators of pedagogical knowledge of evidence-based reading instruction. Answers to each question were checked for accuracy and double-scored by research personnel. For each teacher, the difference in the scores on their pre- and post-test assessments were

calculated. Differences were averaged across teachers within each of the two conditions using the reading curriculum (PD plus coaching and PD only conditions), then the two conditions were compared. The effect sizes were calculated to determine whether teachers provided with coaching support had substantively different scores than those teachers only receiving professional development training (i.e. effect size of .25 or greater). Results are reported in the following chapter.

Student Measures

Screening. A multiple-gating procedure (described previously) was used to identify those students most needing reading intervention in participating classrooms. Teachers were then asked to identify the eight students performing lowest in reading for potential participation. After identifying students based on classroom performance, teachers or members of the research team administered two screening measures.

- 1) Word-reading. To measure word reading ability, a list of fifteen grade-level words separated into three lists was administered. This instrument was used effectively as a screen for first grade students at-risk for reading difficulties in previous studies (e.g., Mathes et al., 2005). The students were shown the first list of five words and asked if they could read any of the words on the list. If the student was able to read one or more of the first five words, s/he was presented with a second list of five words. A third list was presented if the student was able to read one or more words on the second list.
- 2) Texas Primary Reading Inventory Comprehension Task. Teachers administered the Texas Primary Reading Inventory as part of a district-wide assessment and were asked to share students' scores on the listening comprehension task with research personnel. For the comprehension task, students are asked to read two short paragraphs, each followed by

six comprehension questions. One to two of those questions fall into each of the following categories: recalling details, linking details, inferring meaning, and inferring word meaning. If a child is unable to correctly decode three words within the first sentence, the assessor stops administration and reads the paragraph to the child before asking the comprehension questions. If this occurs, the assessment becomes a listening rather than a reading comprehension task.

Pre-test/Post-test Measures.

- 1) Sight Word Fluency Assessment (Denton, 2010). Students were presented with a list of 36 sight words (five per page) and asked to read them as quickly as possible. One point was earned for each word read correctly. The amount of time it took the child to read the list of words was recorded for rate and accuracy calculations.
- 2) Decodable Word Fluency Assessment (Denton, 2010). Students were presented with a list of 40 single-syllable, decodable words in a consonant-vowel-consonant pattern and asked to read them as quickly as possible. One point was earned for each word read correctly. The amount of time it took the child to read the list of words was also recorded for rate and accuracy calculations.
- 3) Decodable Sentence Fluency (Denton, 2010). Students were presented with 26 sentences with 4-6 decodable words each and asked to read as many as possible. One point was earned for each word read correctly within two minutes.
- 4) Oral Reading Fluency Subtest (Good & Kaminski, 2000). On the Oral Reading Fluency subtest of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS), students earned one point for each word of a grade-level reading passage read correctly in one minute. For this study, students read two passages during pre-test and two during post-

test. Correct words read on the two passages during the one-minute allotted time were averaged for each testing period, the mean of which resulted in the child's oral reading fluency score.

- 5) Woodcock-Johnson Tests of Achievement III Letter-Word Identification Subtest (WJ III; Woodcock, McGrew, & Mather, 2001). This subtest required students to identify letters first receptively, by pointing to the letter named by the assessor, then expressively by naming the letter shown on the stimulus. Students were then presented with a list of words and asked to point to certain words and read others. One point was given for each word or letter identified correctly.
- 6) Woodcock Johnson Tests of Achievement III Word Attack Subtest (WJ III; Woodcock, McGrew, & Mather, 2001). On this subtest, students were presented with nonsense words of increasing difficulty. One point was earned for each word read correctly and administration stopped after six consecutive words were read incorrectly.
- 7) WJ III Passage Comprehension Subtest. Students were presented with sentences of increasing length and difficulty, which they read themselves, and were asked to supply a missing word. Correct responses earned one point. Administration ceased after six consecutive incorrect responses.

Intervention Fidelity

Procedural reliability. To measure the extent to which teachers implemented the curriculum as written, fidelity observations were conducted across all teacher participants. Research assistants videotaped intervention sessions three times over the course of the intervention in coached and un-coached teachers' classrooms. Teacher execution was rated using the Reading RULES! Fidelity and Quality of Implementation Instrument and was scored by

trained research personnel using a researcher-made tool, the Fidelity and Quality of Implementation Instrument.

Inter-rater reliability. Observers used videos of study participants from prior years and rated instruction using the fidelity instrument to obtain inter-rater reliability. Videos were viewed and scored by the three trained research personnel, until observers reached 93% agreement across rankings of quality and quantity of intervention components.

CHAPTER 4: Results

Data Analysis

Coaching Fidelity

Percent perfect scores (total number of points awarded divided by the points possible) were calculated for each of the two coaches by granting ten percent for each component rated as present, five percent for each item rated partially present and zero credit for components rated as not present. Each observation had a possible total of one hundred percent. The percent perfect score for the two observations were then averaged to obtain an overall fidelity score for each coach. Coach A received percent perfect scores on her two observations of 75% percent and 90% percent, respectively, for an overall fidelity rating of 83%. When both observations were combined, 90% of the expected components were rated present or partially present for Coach A. Coach B received similar ratings of 80% and 90% after being observed in two classrooms for an average fidelity rating of 85%. Likewise, 85% of the expected components were observed and rated present or partially present during her observed sessions.

Teacher Outcome Measures

Teacher Practice. To obtain a composite score for the Fidelity and Quality of Implementation measure, the teacher's summed ratings were divided by the total possible points resulting in a percent perfect score for each teacher per observation. Composite scores for each observation were then averaged per teacher. The scores of teachers in each of the two groups (PD only and PD plus coaching) were compared and effect sizes calculated for comparison across conditions.

Teacher Knowledge. A one-way analysis of variance (ANOVA) with planned comparisons (PD plus coaching v. PD only, PD plus coaching v. comparison, PD only v.

comparison) was conducted to identify potentially significant differences in post-test outcomes. Effect sizes were also calculated for each of the two groups using the reading curriculum, with effect sizes of .25 or greater considered substantively important.

Student Outcome Measures

Student outcomes on reading measures were grouped into three categories: measures of word reading, measures of reading fluency, and measures of reading comprehension. Separate Analysis of Covariance (ANCOVAs) were used for each reading measure to identify significant differences between students of coached teachers, un-coached teachers and teachers in comparison classrooms. Effect sizes were also calculated and results are reported in the following chapter.

Table 3.
Outcome Measures and Analysis.

	<i>Outcome</i>	<i>Assessment</i>	<i>Development</i>	<i>Statistical Analysis</i>
<i>Teacher Measures</i>	Intervention Fidelity	Fidelity & Quality of Implementation Instrument	Researcher-made	ANOVA & Cohen's <i>d</i>
	Teacher Knowledge	Teacher Knowledge Survey	Researcher-made	ANCOVA & Cohen's <i>d</i>
<i>Student Measures</i>	Word Reading	WJ III Letter-Word ID WJ III Word Attack Sight Word Assessment Decodable Word Reading Assessment	Standardized Test Standardized Test Researcher-made Researcher-made Researcher-made	ANCOVA & Cohen's <i>d</i>
	Reading Fluency	Decodable Sentence Fluency DIBELS – ORF subtest	Standardized Test Standardized Test	ANCOVA & Cohen's <i>d</i>
	Reading Comprehension	WJ III Passage Comprehension	Standardized Test Standardized Test	ANCOVA & Cohen's <i>d</i>

Note: WJ III – Woodcock-Johnson Tests of Achievement III, DIBELS- Dynamic Indicators of Basic Early Literacy Skills; ORF – Oral Reading Fluency, ANOVA – Analysis of Variance, ANCOVA- Analysis of Covariance.

Results

A study was conducted to examine the impact of coaching on the knowledge and practice of teachers and the reading ability of their at-risk first grade students. A Tier II reading curriculum was implemented by 21 teachers in their first grade classrooms in a large urban school district. Of the eight included campuses, five contributed three participating teachers and three campuses contributed two teachers each. Teachers on each campus were randomized into one of three conditions: PD plus coaching, PD only or comparison. As there were unequal numbers of participants by campus, the randomization of the 21 participating teachers yielded six teachers in the PD plus coaching condition, seven in the PD-only condition and eight comparison classrooms. In the eight comparison classrooms, students continued to receive typical instruction in the district selected reading curriculum. The seven classroom teachers in the PD-only condition received professional development training on a curriculum incorporating evidence-based reading instruction and the six teachers randomized to the PD plus coaching condition received regular coaching support in addition to the standard professional development training over the course of the intervention. During the 17 weeks of implementation, students received an average of 72 word-study and comprehension lessons. Teachers randomized to the coaching condition received an average of ten coaching visits over the seventeen weeks of implementation. As coaches worked to be responsive to teachers' needs and work with their availability, some teachers received as few as seven visits while most received ten or twelve. One teacher who demonstrated the need for additional help early in the intervention received coaching support 14 times over the course of the intervention. Videotaped lessons were scored as an indication of fidelity to the curriculum and served as a measure of teacher practice. A survey, consisting of questions related to evidence-based reading instruction was given at pre- and post-test and served as an indicator of

change in teacher knowledge. Four subtests of the Woodcock-Johnson Test of Achievement and three researcher-made assessments were used as measures of student reading ability. Results on outcome measures were analyzed to address the following research questions:

1. When implementing a reading intervention curriculum did teachers who received professional development training and continued coaching support have significantly better outcomes on a measure of implementation fidelity than teachers who received only professional development training?
2. Did students whose teachers received professional development plus coaching have significantly better outcomes in word reading, reading fluency, and reading comprehension than those of teachers who received professional development only or those who implemented typical school instruction? Were the effects of added coaching support substantively important?
3. Did teachers who received professional development training and ongoing coaching support while implementing a reading intervention curriculum have significantly better outcomes on a survey of teacher knowledge than teachers who only received professional development training?

Data Analysis

This study was conducted as part of a larger pilot study; thus it was limited in its scope, sample size and resultant statistical power. Although the random assignment of participants strengthens the model, design limitations suggest the results of this study should be considered an exploratory analysis of the effects of coaching. A more rigorous evaluation should be conducted before results can be considered generalizable and causal connections claimed. Due to the

directional hypotheses presented in each of the three research questions, one-tailed tests of significance were used in the analysis. A Type I error rate of $p = .05$ was set a priori. Analysis of Covariance (ANCOVA) procedures using pre-test scores as the covariates were used to analyze student outcomes and teacher fidelity of implementation scores. A one-way ANCOVA with planned paired comparisons was used to identify statistically significant differences between treatment and comparison conditions (PD only vs. PD plus coaching, PD only vs. comparison, PD plus coaching vs. comparison) to analyze results of the Teacher Knowledge Survey. Effect sizes were also calculated using Cohen's d and results greater than .25 were considered substantively important on all relevant measures (WWC, 2008).

Results are presented in the order of the three research questions addressed by the study: the effects of coaching on teacher knowledge, teacher practice, and student reading outcomes. Statistical outcomes and effect sizes are reported after a discussion of descriptive information is set forth.

Teacher Outcomes

Teacher Knowledge

Tests of Assumptions. Prior to using the Analysis of Covariance to examine post-test difference between conditions, several preliminary analyses were applied to test the assumptions of independence, homogeneity of variances and homogeneity of regression slopes. An analysis of covariance (ANCOVA) was applied to teacher and student data to identify potentially significant differences between groups on post-test scores on the dependent variable. Pre-test scores were used as the covariate to minimize differences in groups prior to the onset of the intervention.

Homogeneity of variance. The homogeneity of variances was assessed using Levene's Test for Equality of Variances and yielded no significant F -statistic ($F=.946, p=.352$), indicating that the assumption was met and equal variances could be assumed.

Test of pre-test differences. An independent samples t-test comparing the means of the two groups was not significant for between group differences. The population means from which the two sample groups were taken was zero ($t= -.698, p=.50$). Differences between pre-test mean scores for coached and un-coached teachers were estimated to be -1.738 , less than one standard deviation apart, a difference that was not statistically significant.

Homogeneity of regression slopes. The assumption of equal slopes was assessed by examining the interaction term (group*pre-test) which was not significant ($F=.146, p=.711$), indicating this assumption was not violated.

Statistical analysis. An analysis of covariance was conducted using the pre-test measure as the covariate. Table 4 reports pre-test and post-test means and standard deviations for the two groups. Adjusted post-test means and standard error terms are also reported.

Table 4.

Teacher Knowledge Survey Pretest, Posttest and Adjusted Means

Condition	Pretest		Posttest		Adjusted Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Coached (n=6)	27.83	3.43	27.50	4.32	28.35	1.15
Un-coached (n=7)	29.57	5.19	32.00	5.23	31.27	1.06

Note: Raw scores were used to calculate means, standard deviations and standard error for the Teacher Knowledge Survey Measure.

The pre-test/post-test correlation was $r = .75$, which indicated a high correlation between pre- and post-test scores. An ANCOVA procedure was selected as an efficient way to compare post-test scores. Results are reported in Table 5.

Table 5.
Teacher Knowledge Survey: Analysis of Covariance Results

ANCOVA	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Covariate	180.46	1	180.46	23.42	.001
Between Groups	26.52	1	26.52	3.44	.047
Within (error)	77.04	10	7.70		
Total	322.92	12			

Note: Type III sum of squares was used.

Results of the ANCOVA were significant for between group differences using a directional alternate hypothesis ($p = .047$, one-tailed). Adjusted post-test mean scores for coached (28.35) and un-coached teachers (31.27) differed by 2.92 points. The direction of the difference favored the teachers who received only the professional development support (no coaching), which was not the hypothesized direction. The pre- and post-test scores for teachers receiving both professional development and coaching support, – 27.83 and 27.50 respectively, was further indication of the lack of impact of coaching on teacher knowledge.

Cohen's d , a measure of effect size, was calculated by subtracting the adjusted post-test mean of the un-coached condition from that of the coached condition and dividing by the pooled standard deviation. As was indicated by the higher post-test means scores and the larger gains of the un-coached teachers, effect sizes favored teachers in the un-coached (PD-only) condition.

Effect sizes for the coached condition were negative ($ES = -.61$), indicating a lack of support for the suggested hypothesis, that teachers in the PD plus coaching condition would significantly outperform those teachers who only received professional development support. Overall, results indicated greater growth on the knowledge measure for the un-coached teachers in this exploratory analysis. It should be noted that differences in both groups were not statistically different at pre-test (less than two points) or post-test (less than three raw score points).

Teacher Practice

Test of Assumptions

Teachers in both conditions were videotaped at three points over the course of the intervention. Fidelity scores were calculated by dividing the points scored on the fidelity of implementation instrument by the total points possible for the completed components. One teacher in the un-coached condition was unable to be filmed for the third taping session and thus is missing a data point. The statistical software excluded this case from the analysis. Review of the teacher's scores revealed that her first video score was higher than the group mean and the second score was lower than that of the rest of the un-coached group. These differences make it difficult to predict how a third score would have impacted the results.

Independent samples. A t-test for independent samples was significant at the first time period and not significant at each of the two subsequent time periods. The results for video one ($p=.052$), video two ($p=.476$) and video three ($p=.178$) indicated there is not evidence that fidelity scores between the two groups were sufficiently different to meet this hypothesis.

Sphericity. Because these data were collected from repeated measures, testing the sphericity assumption is critical. Mauchly’s test of sphericity was significant ($p=.001$) indicating the variances of differences between levels of related variables were not equal. The test for sphericity is one way of examining the equality variances and covariances and is closely related to tests for homogeneity of variances. Because this assumption was not met, the Huynh-Feldt correction was applied in the analysis.

Table 6.

Teacher Implementation Fidelity

Condition	Video 1		Video 2		Video 3	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Coached (n=6)	95.47	3.38	95.88	4.30	93.52	5.58
Un-coached (n=6)	88.46	9.24	93.71	6.06	89.40	10.55

Descriptive information for the two groups is reported in Table 6. The three videotaping time periods represent three levels of the independent variable for this analysis. For each of the three levels of the independent variable, teachers in the professional development plus coaching condition exhibited higher mean scores and smaller standard deviations than those teachers who only received professional development. Both groups scored highest on their second fidelity observation and scores decreased slightly on their final observation. Coached teachers’ scores were higher over time than those teachers who only received professional development.

A Repeated Measures Analysis of Variance (ANOVA) was used to compare the two groups across the three levels of the independent variable. As analyzing each observation period alone

was unlikely to reveal significant differences between groups, the use of a Repeated Measures ANOVA allowed for exploration of pooled observations, creating a more stable estimate and resulting in higher observed power for the analysis.

Analysis of Intervention Effects Results indicated the professional development plus coaching condition's overall mean (pooled across time) was significantly higher than the group of teachers who only received professional development training when a one-tailed criterion was used (see Table 7). The test of differences for the three time intervals (group by time interaction) was not significant (see Table 8).

Table 7.

ANOVA Results: Between Teachers Effects

Repeated Measures ANOVA	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Groups	197.40	1	197.40	3.5	.046
Within Groups (error a)	546.05	10	56.40		

Table 8.

ANOVA Results: Within Teachers Effects

Repeated Measures ANOVA	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Time	52.20	2(1.275)	90.94	1.182	NS
Group * Time	19.82	2(1.275)	15.54	.202	NS
Time * Within Groups (error b)	980.94	20(12.75)	76.93		

The measurable effects of coaching on the knowledge of teachers of at-risk first grade students were not significantly different than those of in an un-coached condition, but results did

support coaching as an agent of change in teacher practice. On a measure of teacher knowledge, both groups exhibited high initial scores and neither group made significant gains at post-test. Although both groups of teachers performed high on initial measures of implementation fidelity, teachers who received coaching support after attending professional development trainings did show significant differences in rating scores when compared to teachers who only received professional development. While coaching support did not appear to impact teacher knowledge over the course of the intervention, it does appear to have impacted and improved teacher practice over the same period of time. Teachers in both conditions received a total of four days of professional development training; those with coaching support in addition to the trainings scored significantly higher over the course of the intervention on a measure of implementation fidelity.

Student Outcomes

Test of Assumptions To examine the impact of coaching teachers on the reading outcomes of first-grade students, pre- and post-test data were collected on a battery of researcher-made and standardized assessments. To continue examining the data on the teacher level, student scores in each classroom were averaged on each outcome measure prior to running the analyses. High pre- and post-test correlations indicated an Analysis of Covariance (ANCOVA) was appropriate and are reported in Table 9. The ANCOVA was then used to examine student growth while minimizing pretest differences between groups. Results are divided into measures of word reading, reading fluency and reading comprehension.

Table 9.

Pre-test-Posttest Correlations for Student Outcome Measures

	Letter-Word Identification	Word Attack	Sight Word Fluency	Decodable Word Fluency	Decodable Sentence Fluency	Oral Reading Fluency	Passage Comprehension
Pre-post correlation	.811	.868	.651	.842	.682	.766	.467

Word Reading Measures

Table 10.

Student Word Reading Outcomes: Homogeneity of Variances

	Letter Word Identification		Word Attack		Sight Word Assessment		Decodable Word Assessment	
	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>
Levene's Test	.375	.693	1.179	2.07	3.28	.061	.473	.630

Homogeneity of variance. Levene's Test was used to test the assumption of homogeneity of variances. The F-statistic for each of the four word reading outcome measures was not significant: Woodcock-Johnson Letter-Word Identification subtest ($F = .375, p = .693$), Woodcock-Johnson Word Attack subtest ($F = 1.179, p = .207$), Sight Word Assessment ($F = 3.28, p = .061$) and the Decodable Word Assessment ($F = .473, p = .630$). The lack of significance indicates the assumption of equality of variances was met (see Table 10).

Pre-test comparisons. A One-way Analysis of Variance (ANOVA) was conducted to compare means at pre-test on the four word reading measures and determine whether the means of group were different at pre-test. Results were not significant between the means of the groups at pre-test for the four dependent variables (Letter-word identification, $p=.453$, Word Attack, $p=.614$, Sight Word Assessment, $p=.356$, Decodable Word Assessment, $p=.812$).

Homogeneity of regression slopes. Effects for the group by measure interaction were not significant for each of the four reading outcomes: letter-word identification ($p=.189$), word attack ($p=.154$), sight words ($p=.734$) and the decodable word assessment ($p=.363$) indicating that homogeneity of regression slopes could be assumed.

Table 11.

Word Reading: Pre-test, Post-test and Adjusted Post-test Mean Scores

<i>Measure</i>	<i>Pre-test</i>		<i>Post-test</i>		<i>Adjusted Post-test</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
WJ Letter-Word Identification						
Coached (n=6)	389.51	8.95	421.30	3.63	414.67	4.26
Un-coached (n=7)	377.89	11.64	416.76	18.79	422.18	3.92
Comparison (n=8)	382.90	22.46	406.21	25.85	406.88	3.60
WJ Word Attack						
Coached (n=6)	440.54	15.77	464.34	9.83	459.75	2.92
Un-coached (n=7)	433.05	16.94	459.87	13.23	461.20	2.66
Comparison (n=8)	431.83	17.88	449.71	19.07	451.99	2.50

Table 11 (continued)

Sight Word Assessment							
Coached (n=6)	8.44	4.08	43.42	4.50	44.49	3.29	
Un-coached (n=7)	6.10	3.26	32.61	12.75	37.07	3.12	
Comparison (n=8)	12.4	12.60	27.94	19.45	23.25	2.94	
Decodable Word Assessment							
Coached (n=6)	3.46	2.03	13.91	2.63	13.82	1.25	
Un-coached (n=7)	2.74	2.44	11.90	5.10	13.00	1.17	
Comparison (n=8)	3.95	5.07	10.21	9.24	9.33	1.09	

Note: Raw scores were used to compute means, standard deviations and standard error on researcher-made measures and W-scores were used for Woodcock-Johnson measures.

Analysis of Intervention Effects. To examine the impact of coaching teachers on their students' decoding and comprehension abilities, an Analysis of Covariance was conducted using the four word reading measures (Letter-Word ID, Word Attack, Sight Words, Decodable Words) as dependent variables. An ANCOVA analysis was used to reduce the error variance and consequently improve the efficiency and power of post-test means by co-varying out group differences on pre-test measures. Due to the small sample size, using a Bonferroni correction would have reduced the power of the analysis, setting a more stringent Type I error rate, which was not appropriate for this type of exploratory study. Results for each of the four dependent word-reading variables are presented in Table 11.

Results of the omnibus test were significant for differences between the three groups (see Table 13). Further analysis indicated these differences were found between the intervention and comparison groups only and no statistically significant difference between coached and un-coached groups on of the four student outcome measures were identified. Post-hoc pairwise

comparisons demonstrated students in the coached condition on the two proximal word reading measures (sight words and decodable words) yielded higher levels of significance over students in the comparison condition than did those students of un-coached teachers (Table 12). Still, results clearly favored the students of un-coached teachers over those in the coached condition when measured against the comparison condition, as they significantly outperformed comparison students on all four measures of word reading.

Table 12.

Post-hoc Pairwise Comparisons for Adjusted Post-test Means.

	Coached vs. Un-coached		Coached vs. Comparison		Un-coached vs. Comparison	
	Mean Difference	<i>p</i>	Mean Difference	<i>p</i>	Mean Difference	<i>P</i>
WJ Letter-Word Identification	-7.507	.222	7.791	.181	15.297	.010*
WJ Word Attack	-1.438	.722	7.764	.062	9.203	.022*
Sight Word Assessment	7.419	.118	21.239	.000*	13.820	.006*
Decodable Word Assessment	.821	.638	4.487	.015*	3.665	.036*

Table 13.

Results of Analysis of Covariance for Word Reading Measures

WJ Letter Word Identification

ANCOVA	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Covariate	5041.20	1	5041.20	48.68	.001
Between Groups	862.33	2	431.16	4.16	.017
Within (error)	1760.53	17	103.56		
Total	7664.05	20			

WJ Word Attack

ANCOVA	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Covariate	3682.42	1	3682.42	74.73	.001
Between Groups	326.95	2	6181.47	3.68	.024
Within (error)	837.70	17	49.28		
Total	4883.06	20			

Sight Word Assessment

ANCOVA	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Covariate	1946.22	1	1946.22	30.09	.001
Between Groups	1542.51	2	6771.25	11.94	.0005
Within (error)	1099.60	17	64.68		
Total	4588.32	20			

Table 11 (continued)

Decodable Word Assessment

ANCOVA	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Covariate	592.62	1	592.62	62.87	.001
Between Groups	82.16	2	641.08	4.36	.015
Within (error)	60.24	17	9.43		
Total	835.02	20			

Note: Type I Sum of Squares was used.

Effect sizes. Estimates of effect size, reported in Table 14, make the differences between conditions more clear. Effect sizes were calculated by subtracting the adjusted mean of the professional development plus coaching or comparison condition from the adjusted mean of the variable of interest (coached or un-coached) and dividing by the pooled standard deviations of the posttest scores for the groups being compared. An effect size of .25 is regarded as substantively important for the purpose of this study (WWC, 2008).

Table 14.

Student Outcomes: Pre-test to Post-test Adjusted Mean Difference Effect Sizes

Group	Coached vs. Un-coached	Coached vs. Comparison	Un-coached vs. Comparison
Letter-Word ID	-.55	.42	.68
Word Attack	-.12	.51	.56
Sight Words	.78	1.50	.84
Decodable Words	.20	.66	.49

These estimates further illustrate the differences between conditions and again indicate positive effects for coached classrooms over those in the professional development only condition

for the two researcher-made measures, but highlight a negative effect on the standardized Woodcock-Johnson subtests (Letter-word ID, $ES=-.55$, Word Attack, $ES= -.12$). While it is more difficult to make significant gains on standardized measures over a short period of time, the fact that strong positive effects were made by the un-coached and coached conditions demonstrate that positive measurable gains are attainable in a seventeen-week timespan, yet not successfully attained by those in the comparison condition.

One potential reason for the lack of significant improvement of coached classrooms over un-coached participants may be the strength of the reading intervention itself. As observed, both students in the coached and un-coached conditions made considerable gains over those in the comparison group, and made similar post-test scores on each of the four measures, with post-test adjusted means for the two conditions being less than one raw score point apart on a measure of decodable word reading. Post-test scores on the Woodcock-Johnson Word Attack subtest for coached and un-coached conditions were only 1.45 points apart when adjusted for differences in pre-test. Although additive effects on two measures (sight words and decodable words) for those in the coached condition were present when compared to peers in the comparison condition, it is apparent that students receiving the intervention curriculum made substantial gains on proximal measures whether their teachers were coached or un-coached. All students in the un-coached condition made statistically significant gains over their comparison counterparts.

Reading Fluency Measures

Test of Assumptions. Identical procedures were used to run analyses for student measures of reading fluency. Levene's test yielded non-significant F -statistics for both the Oral Reading Fluency measure ($F=.624$, $p=.547$) and the Decodable Sentence Fluency measure ($F=3.406$, $p=.056$), thus equal variances were assumed (see Table 15). A One-way Analysis of

Variance was used to compare means between groups at pre-test and no statistical differences were observed for either of the two fluency measures. Pre-, post- and adjusted post-test means are reported in Table 13. Homogeneity of regression slopes was examined by looking for significant interactions between pre-test and condition. Results for the fluency measures (Oral Reading Fluency, $F=.404$, $p=.675$; Decodable Sentence Fluency, $F=.591$, $p=.566$) were not significant, indicating this assumption was met.

Table 15.

Reading Fluency: Homogeneity of Variances

	DIBELS Oral Reading Fluency		Decodable Sentence Fluency	
	<i>F</i>	<i>p</i>	<i>F</i>	<i>P</i>
Levene's Test	.624	.547	3.406	.056

Examination of pre- and post-test scores on both measures of reading fluency indicate the comparison condition students read at least two more correct words per minute on the Decodable Sentence Fluency measure and also read more words correctly on the Oral Reading Fluency measure than those in the coached and un-coached condition. As anticipated, those in the comparison group scored lowest on the same measures at post-test. In contrast, students in the coached condition exhibited a mean score between that of students in the comparison and un-coached conditions at pre-test but yielded the highest means at post-test on both measures.

Students of coached teachers made greater gains at post-test on the Decodable Sentence Fluency measures than those in the un-coached and comparison conditions. This effect did not

hold true when adjusted post-test means were used for comparison across groups. After co-varying for pre-test scores, those in the un-coached condition yielded the greatest gains. On the DIBELS Oral Reading Fluency measure, gains were largest for students in the coached condition whether adjusted or un-adjusted means were used.

Table 16.

Reading Fluency: Pre-test, Post-test and Adjusted Post-test Mean Scores

<i>Measure</i>	<i>Pre-test</i>		<i>Post-test</i>		<i>Adjusted Post-test</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
<i>Decodable Sentence Fluency</i>						
Coached (n=6)	7.07	2.72	31.83	2.12	32.21	2.43
Un-coached (n=7)	4.94	1.83	27.35	9.75	31.09	2.32
Comparison (n=8)	9.58	9.20	22.43	15.19	18.87	2.17
<i>DIBELS Oral Reading Fluency</i>						
Coached (n=6)	3.37	2.06	22.02	5.22	23.00	1.98
Un-coached (n=7)	2.22	1.48	15.66	6.34	18.54	1.87
Comparison (n=8)	3.97	4.81	16.28	12.74	13.02	1.77

Note: Raw scores were used to compute means, standard deviations and standard error on researcher-made and DIBELS measures.

Analysis of Intervention Effects. Pre-test scores for the two measures of reading fluency were used as covariates in the Analysis of Covariance to examine effects of the coaching intervention on the dependent variables. A significance level was set a priori at .05 and no corrections for family-wise error rate were applied due to the small sample size.

The results of the omnibus test (Table 17) were significant for differences between the three conditions. One-tailed post-hoc analyses were conducted by pairing conditions to identify potentially significant differences between groups. Results indicated no statistically significant differences between the coached (PD+ coaching) and un-coached (PD-only) groups on either the Decodable Sentence Fluency or Oral Reading Fluency measures. These pairwise comparisons mirrored the results of the ANCOVA, which yielded no significant difference between student post-test scores in the coached and un-coached conditions. What did emerge in this observation of findings was an additive effect for those in the coached condition when paired with the comparison students. Students in classrooms of coached teachers did demonstrate significant differences on both measures of reading fluency when compared with students in the comparison condition. Students of un-coached teachers only showed significant differences at post-test on one measure of reading fluency (decodable sentences) when paired with comparison students.

Table 17.

Results of Analysis of Covariance for Reading Fluency Measures

ANCOVA: Decodable Sentences					
	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Covariate	1171.93	1	1171.93		.001
Between Groups	739.10	2	369.55	26369.55	.0005
Within (error)	603.70	17	35.51		
Total	2514.73	20			

ANCOVA: Oral Reading Fluency					
	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Covariate	949.60	1	949.60	40.75	.001
Between Groups	326.25	2	163.13	7.00	.003
Within (error)	396.14	17	23.02		
Total	1671.99	20			

Table 18.

Reading Fluency Outcomes: Post-hoc Pairwise Comparisons

	Coached vs. Un-coached		Coached vs. Comparison		Un-coached vs. Comparison	
	Mean Difference	<i>p</i>	Mean Difference	<i>p</i>	Mean Difference	<i>p</i>
Decodable Sentence Fluency	1.12	.372	13.34	.0005*	12.22	.001*
DIBELS Oral Reading Fluency	4.46	.0585	9.99	.001*	5.52	.027

Effect sizes. Estimates of effect size also indicated the smallest effects for the coached versus un-coached groups. Effect sizes for those in the coached condition versus the comparison conditions were consistently stronger than those of the un-coached versus comparisons on both measures of reading fluency. Effect sizes are reported in Table 18.

Table 19.

Reading Fluency Outcomes: Adjusted Mean Difference Effect Sizes

Group	Coached vs. Un-coached	Coached vs. Comparison	Un-coached vs. Comparison
Decodable Sentence Fluency	.16	1.23	.96
DIBELS Oral Reading Fluency	.77	1.03	.55

Reading Comprehension Outcomes

To measure effects of coaching and a Tier II reading intervention on students' reading comprehension abilities, the Woodcock-Johnson Passage Comprehension subtest was administered as part of the pre-/post-test assessment battery. The assumptions of homogeneity of variances was examined using Levene's Test ($F=1.37, p=.279$). Resulting p-values indicated the homogeneity of variances was met. Homogeneity of regression slopes was assumed after examining the interaction between pre-test scores and condition, which was not found to be significant ($p=.901$).

Descriptive Data. Observation of pre- and post-test means indicated students in the comparison condition began with the highest W-scores at pre-test. After the implementation of both the reading and coaching interventions, examination of post-test

and adjusted post-test scores indicated students in the coached condition demonstrated the greatest gains in passage comprehension. Table 20 reports descriptive data.

Table 20.

Reading Comprehension: Pre-test, Post-test and Adjusted Post-test Mean Scores

<i>Measure</i>	Pre-test		Post-test		Adjusted Post-test	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
WJ Passage Comprehension						
Coached (n=6)	421.25	7.74	451.03	7.57	450.76	4.72
Un-coached (n=7)	417.33	7.63	444.50	11.75	446.82	4.48
Comparison (n=8)	423.62	11.72	443.90	13.33	436.20	4.16

Note: W- scores were used to calculate means, standard deviations and standard error for Woodcock Johnson measures.

Analysis of Intervention Effects. The Analysis of Covariance yielded no statistically significant differences between students of coached and un-coached participants on the comprehension variable (see Table 21). Post-hoc analysis pairing each of the three conditions indicated the only significant difference were W-scores of students in the coached condition when paired with the comparison group receiving typical instruction. Again, this may demonstrate an additive effect of coaching as scores of students in the un-coached condition were not significantly different than those in the comparison (see Table 22).

Table 21.

Results of Analysis of Covariance for Reading Comprehension

ANCOVA	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P</i>
Covariate	495.35	1	495.35	3.71	.071
Between Groups	787.40	2	6393.70	2.91	.080
Within (error)	2269.23	17	33.48		
Total	3551.98	20			

Table 22.

Student Reading Comprehension Outcomes: Post-hoc Pairwise Comparisons for Adjusted Means

	Coached vs. Un-coached		Coached vs. Comparison		Un-coached vs. Comparison	
	Mean Difference	<i>p</i>	Mean Difference	<i>p</i>	Mean Difference	<i>p</i>
WJ Passage Comprehension	3.94	.272	14.56	.165*	10.62	.0535

Effect sizes. Once again effect sizes were calculated by subtracting group means and dividing by the pooled standard deviation. Results are reported in Table 23. Effect sizes for each of the three paired comparisons (coached v. un-coached, coached v. comparison, un-coached v. comparison) were each large positive effects with the strongest effect indicated for the coached versus comparison condition.

Table 23.

Reading Comprehension Outcomes: Adjusted Mean Difference Effect Sizes

Group	Coached vs. Un-coached	Coached vs. Comparison	Un-coached vs. Comparison
WJ Passage Comprehension	.40	1.34	.85

Summary. Analysis of student outcome measures yielded no significant effects on measures of reading achievement for students in coached versus un-coached (PD-only) conditions. Results of four word-reading measures indicated that all students in the un-coached (PD-only) condition outperformed those in the comparison group, whereas students in the coached (PD plus coaching) condition only demonstrated significantly higher post-test scores on the two word-reading measures proximal to the intervention: sight words and decodable words. On the two measures of reading fluency, students of coached teachers outperformed the control group students on both decodable sentence fluency and DIBELS oral reading fluency measures. Students of un-coached teachers had significantly different post-test scores on one measure: decodable sentence reading fluency. Post-hoc analysis of scores on the passage comprehension subtest of the Woodcock-Johnson Test of Achievement indicated only the coached versus comparison conditions exhibited statistically significant differences in favor of the coached condition, while all three paired comparisons demonstrated strong positive effect sizes.

CHAPTER FIVE: Discussion

The effective use of evidence-based instructional practices for the teaching of reading has been consistently shown to improve student reading outcomes over time (Kretlow & Bartholomew, 2010). Differences in student achievement are often contributed to what the literature refers to as a “teacher effect,” as teachers with working knowledge of the phonological, morphological and linguistic underpinnings of the reading process have been shown to be most effective at improving student reading outcomes (Carlisle, Kelcey, Rowan & Phelps, 2011; Moats, 1994).

In addition to an adequate knowledge of reading processes, educators must be able to review current research and select practices that would be most beneficial for struggling readers within their classrooms. Further, instructors employing evidence-based instructional practices with their at-risk readers need to be able to evaluate their effectiveness, collect and interpret student performance data, and gauge students’ responses to intervention. Often the skills teachers need in order to assess, revise, and adapt their own practice are not acquired prior to their experience in the classroom; thus, new skills must be learned through professional development seminars or trainings.

In a recent review of studies, coaching was demonstrated to improve the fidelity with which teachers applied newly acquired instructional skills (Kretlow & Bartholomew, 2010), yet a convergence of evidence does not yet support the practice. Fewer studies have reported student academic outcomes correlated with coaching support and, of those, results were often only reported for specific skills (e.g. alliteration) or subgroups of students (e.g. English Language Learners) (Cusumano, 2006; Matsumara, 2010). The current study, while exploratory, examined

the effects of coaching on all three variables: teacher knowledge, teacher practice (as measured by implementation fidelity), and student reading achievement outcomes.

This study included twenty-one classroom teachers grouped by campus and then randomly assigned to a coached (n=6), un-coached (n=7), or control (n=8) condition. Both coached and un-coached teachers received professional development training in a newly developed Tier II reading intervention and were asked to implement the curricula with their five lowest-performing students. For the coached group, coaching support was provided on a bi-weekly basis, for approximately one hour per week, with a twenty to thirty minute follow-up discussion per observation. Scores on a measure of teacher knowledge yielded no significant differences for coached participants over their un-coached counterparts. Teacher practice ratings on an implementation fidelity instrument were significantly higher for teachers in the coaching plus professional development condition than those in the professional development only condition. Results on measures of student reading achievement did not yield significantly higher outcomes for students of coached teachers over students of un-coached participants, yet in a few cases, promise of additive effects was demonstrated.

Findings by Research Question

Teacher Knowledge

The first hypothesis addressed in the study was the belief that coached teachers would demonstrate significant improvements in knowledge of evidence-based practices over their un-coached counterparts. An analysis of pre- and post-test scores on a researcher-made measure of teacher knowledge indicated no significant differences between coached and un-coached teachers. At post-test, teachers who did not receive coaching had higher scores on the knowledge

survey than those who received coaching support. Teachers who received coaching support demonstrated a moderate negative effect size (-.61), in opposition to the hypothesis.

In post-intervention interviews with participants, both coached and un-coached teachers self-reported a dramatic increase in their understanding and ability to use evidence-based reading interventions; these results were not reflected in posttest outcome measures. Findings such as these are not unique. Research suggests that perceived growth does not always translate into measurable data and has led some investigators to reflect on the sensitivity and validity of the measures we use to examine change in teacher knowledge (Carlisle, Cortina & Katz, 2011).

Another plausible explanation for the lack of growth in teacher knowledge may be the duration of instructional support provided to the teachers themselves. While students in the study received seventeen weeks of intervention, teachers received only bi-weekly coaching support. With coaching visits lasting 60-90 minutes per visit on average, teachers received roughly 7 to twenty hours of “instructional time” through consultative conversations with coaches. This is far less than the two years of professional development support suggested as necessary to make measurable growth in educators (Moats, 1994).

Studies which have increased the duration and rigor of instruction provided to educators have yielded similar findings and suggested a third possible explanation for their lack of significant findings (Neuman & Wright, 2010). Neuman and Cunningham (2009) provided teachers in their study with 45 contact hours in a community college course on early childhood literacy instructional practices, yet post-test scores also resulted in no significant effects on a measure of teacher knowledge. Further analysis of their pre-test data reflected a similar pattern to the current study, as teachers in both conditions began the study with a high-level of knowledge

of the content measured. Thus, in both the current study and that of Neuman and Cunningham (2009), teachers may have had little room for growth.

A final consideration is mirrored in the findings of Armstrong, Cusuamano, Todd and Cohen (2008). In studies where new, research-based interventions are employed, teachers often benefit. Growth, although incremental, appears to be a product of the strong instructional materials themselves; coaching support has not proven to produce statistically significant gains in knowledge enough to distinguish between un-coached implementers. In their study, although none of the teachers in either condition achieved eighty percent mastery on the nineteen-item survey of teacher knowledge at pre-test, participants in both conditions made significant gains at post-test (Armstrong, Cusuamano, Todd & Cohen, 2008). Still, post-test scores could not differentiate between conditions; coached teachers' mean post-test score (13.70) was not significantly different from those of the un-coached comparison (13.63).

Teacher Practice

The second hypothesis tested by this study was that teachers receiving regular coaching support would demonstrate significantly greater improvements in observed practice than their un-coached counterparts. In the current study, evidence-based practices were built into a Tier II reading intervention and changes in teacher practice were measured using an implementation fidelity instrument. Fidelity data was collected at three points during the intervention and percent perfect scores were calculated for each session per teacher, as described in chapter three. Using a one-tailed, directional hypothesis, ANOVA repeated measures analysis indicated significant differences between groups, favoring those in the coached condition.

The significance indicated in the one-tailed test is supported by previous research which shows that coaching support does produce measurable gains in teacher implementation of reading practices and associated changes in instructional behavior. In a review of 13 studies, Kretlow and Bartholomew (2010) reported moderate effect sizes of increased implementation fidelity of evidence-based practices when coaching support was provided. Similarly, in 2010, Sailors and Price reported on two studies in which teachers made significant gains, after coaching, in providing increased opportunities for students to engage in comprehension activities as well as providing more structured explanations of the reading strategies being taught. Although multiple studies report changes in teacher practice after coaching support is provided, studies often report change in more general teacher behavior or in the learning environment, frequency and variation of literacy activities, access to print materials, student engagement and praise, rather than changes specific to reading instruction (Neuman & Cunningham, 2009; Sailors & Price, 2010).

In the current study, teachers attended two, full-day professional development trainings, two follow-up professional development trainings and received bi-weekly coaching visits. In contrast, Carlisle, Cortina and Katz (2011) provided their participants with nine professional development seminars (twenty-seven instructional hours) and coaches attended or led grade-level meetings on a weekly basis. This increased amount of instructional hours as well as frequency of coaching visits may have produced more significant outcomes for the participants.

Of additional importance in relation to fidelity outcomes in the current study are teachers' improved scores for both conditions on the second fidelity observation, which was conducted shortly after the second professional development training. For this training, teachers were excused from their regular teaching duties and visited another campus to observe two teachers implementing the intervention. Research personnel and coaches then discussed both their

observations and the suggested ways that they might include some of the strategies they observed. The higher scores during second fidelity rating period may well have been the consequence of this type of professional development experience.

The number of lessons taught by a teacher over the course of the intervention may also have played a role in their quality of implementation. Coaches' anecdotal notes indicated teachers who taught a greater number of lessons overall, or who did so more frequently, became more familiar with the routines set forth in the intervention. They were therefore more likely to manipulate materials with ease, increase instructional time and engage students than those who implemented the curriculum less frequently. With these teachers, coaches were more apt to focus on refining instructional components, reviewing data and co-planning than with teachers who were struggling with lesson preparation, organization of materials and behavior management.

Student Reading Achievement

Research on the impact of teacher coaching on student reading achievement outcomes has yielded mixed results, similar to those reported in the current study. The current study used a variety of researcher-made and standardized assessments to break down student performance into three areas: word reading, reading fluency, and reading comprehension. Although students in the coached condition displayed impressive growth on each of the seven measures of reading achievement, they did not show significant gains over students receiving the intervention curriculum from teachers who only received professional development training.

Additive Effects. Differences between the two treatment conditions (coached and un-coached) were also compared to the comparison condition, where students received typical reading instruction. Students in the professional development plus coaching condition

demonstrated significant effects over students in the comparison group on five of the seven student outcome measures. Many of these positive effects were shared by student participants in the professional development only condition when examining differences between comparison (no treatment) students. Yet, on three measures -- decodable sentence assessment, DIBELS Oral Reading Fluency Assessment and WJ Passage Comprehension subtest -- results showed a possible additive effect for students in the coached condition. On these measures, students of coached teachers had significantly higher post-test scores than those in the comparison condition while results from those in the un-coached and comparison conditions were not significant.

Similar studies report that coaching has little to no measureable impact on student reading outcomes (Marsh, et al. 2008, Garet, et al. 2008). In a study of forty-one kindergarten teachers receiving coaching, Cusumano, Armstrong, and Todd (2006) also yielded no significant results in student skill attainment. Yet, these findings conflict with a larger body of related research which suggests coaching support has positive measurable impact on student achievement.

Hsieh, Hemmeter, McCollum, and Ostrosky (2009) used a similar coaching model to the one implemented in the current study. Teachers were taught twenty teaching strategies, grouped into three clusters, and were then supported by two coaches over the course of a six-week intervention period. A similar observation-and-feedback coaching model was used and teachers were provided with varied number of coaching visits determined by need. In contrast to the current study, their teachers were supported on average two to three times per week. Results on outcome measures of picture naming, alliteration, rhyming and print knowledge indicated significantly higher post-test scores on all four assessments when in the coached condition. The authors reported “measures of the children’s learning provide preliminary evidence that even a

short period of coaching may be associated with relatively rapid changes in children's emergent literacy skills" (p. 243)

The idea that low-performing students may receive the most benefit from coached teachers may be another avenue for further exploration and was suggested by anecdotal data in the current study. Biancarosa, Bryk, and Dexter (2010) conducted a longitudinal study that found that students with lower initial literacy scores, whose teachers took part in a literacy collaborative with a coaching component, learned at a faster rate than those who entered with higher literacy abilities and that these improvements increased over time (Biancarosa, Bryk & Dexter, 2010). At least one recent multi-site study found English Language Learners made significant gains on state standardized assessments when their teachers received coaching support and received greater benefit from such support than their peers in the same classrooms (Matusmara, Garnier, Correnti, Jimmer & Bickel, 2010).

Summary and Analysis

Findings from this study reflect those of similar studies of coaching which have yielded mixed results (Matsumara, et al. 2010). Prior research has indicated stronger support for the use of coaching to increase teacher knowledge and practice (Knight, 2009). Studies that have presented positive correlations between the coaching of teachers and student reading outcomes, though less frequently reported, have demonstrated significance for improvements in specific reading sub-skills (e.g. comprehension, phonemic awareness) and have most often been conducted with preschool age children. Often, these interventions were applied to classrooms in their entirety, rather than targeted for students at-risk for reading failure as were those in this study.

Differences in the outcomes presented in this study and those that have previously provided stronger support for the practice of coaching stem from several potential factors. The sensitivity of measures of teacher knowledge and practice are of concern for the authors of each of the supporting studies, as is the case with the current study. Differences in the type and duration of coaching activities employed and the robust nature of the Tier II intervention curriculum are also a point for further discussion.

Measurement. The lack of adequate and sensitive measures to assess teacher content and pedagogical knowledge is not a new issue for researchers. Phelps and Schilling (2004) used exploratory factor analysis to determine which of their 261 questions, field-tested with more than 1,500 California school teachers, would result in valid and reliable measures of two types of teacher content knowledge. Using rigorous psychometric analysis, their study resulted in a bank of questions that could effectively measure teacher knowledge in the areas of word analysis and comprehension. Unfortunately, the qualifying questions were not able to reliably assess a wide range of teacher knowledge; in other words, items that loaded on the word analysis factor were relatively easy items and useful only for teachers who knew little about the teaching of reading, while items that loaded on the comprehension factor were relatively difficult and best used to distinguish between those teachers and others who already had a strong foundation in reading comprehension instruction. This analysis illustrates the relative difficulty in creating valid and reliable measures that are sensitive to a wide range of teacher knowledge.

Coaching activities. Activities most often used in the coaching model presented in the current study reflect those most often found in the literature: a pattern of observation, scaffolding and feedback provided to supported teachers.

A review of coaching logs indicated that of the 78 coaching visits provided during the seventeen-week intervention, 60 or more included a pattern of observation and a debriefing discussion to determine what went well and what could be improved about the lesson as a whole. Only eight of these sessions reviewed and analyzed student assessment data and eleven included modeling. Even fewer visits, six in each case, included side-by-side teaching or co-planning. Anecdotal data indicated coaches spent a majority of their time in classrooms supporting teachers' efforts at classroom management or handling logistical questions and concerns regarding the curriculum, practices not aimed at improving practice or pedagogical knowledge.

Duration and Intervention Strength. The length of time for which teachers receive coaching support has been considered a limitation in the current study, yet coaching has been shown to have measurable impact on students whose teachers received as few as six collaborative coaching visits and to have additive benefits when teachers were consistently coached over a period of several years (Biancarosa, Bryk, & Dexter, 2010; Sailors & Price, 2010). Instead, the robust effects of the Tier II intervention itself may actually have been the largest catalyst for change in the current observation of implemented practices. As teachers in both treatment conditions were new to the use of the curriculum, and as students in the intervention made significant measurable gains when compared to their peers receiving typical instruction, it may be important to consider the steep learning curve all teachers had in implementing the intervention. Under these circumstances, coaching may not have had as powerful an impact on this novel implementation as it would have otherwise.

Implications for Practice

This study adds to the body of literature indicating mixed results on the effects of

coaching support provided to classroom teachers. Whereas coaching continues to be a widely accepted and commonly used practice with high social validity ratings, a lack of consensus exists as to its effectiveness. Thus, administrators may be promoting a practice without a solid body of empirical evidence supporting its impact. Results of this exploratory study should encourage further examination of coaching and other forms of sustained professional development to identify the methods and sustainable tools that best promote improved educational outcomes for educators and their students.

Differences in the definition and application of coaching constructs make it difficult to compare outcomes across studies and create a cohesive body of literature supporting the practice. As researchers use different types or components of coaching across studies, it becomes difficult to clearly indicate causal connections between coaching practices and desired outcomes. The lack of adequate description of coaching components within studies also creates barriers for those seeking to replicate or generalize findings. Thus, the findings presented here may reflect a broader lack of clearly outlined and applied evidence-based application coaching practices across the field rather than a failure of coaching support to produce desired outcomes within one particular study. Although the reading intervention lessons included evidence-based practices such as explicitly and systematically taught skills, multiple opportunities for practice and daily supported interaction with text, the professional development and coaching activities were used to support implementation of the intervention rather than disseminate information and research supporting said instructional techniques. Instead of teaching the teachers the reasons why we were instructing students a certain way, coaches more often than not found themselves scaffolding for successful inclusion of lesson components, providing corrective feedback on activities and troubleshooting behavioral and logistical concerns.

Of greater import and potentially greater difficulty may be the delineation of the practice of coaching itself. By its nature, coaching is a responsive practice rather than a standardized or rigid set of skills or behaviors. Commonly without systematic or sequential steps, the most effective form of coaching has been found to be dynamic and responsive in nature. This pattern of coaches responding to teachers who are responding to students creates a cycle of instruction where both coach and teacher meet their respective “student” at their current ability level and provide remediation, instruction, modeling and collaborative (co-teaching, co-planning) support for correct responses and implementation.

To promote greater impact on the knowledge and practice of teachers, coaches must shift from the role of passive observer to active instructor. Rather than responding to inadequate instruction, coaches should provide instruction, modeling and rationale for the use of evidence-based practices in the classroom. The role of coach is to create greater capacity within an educator to become better evaluators of information and improve their ability to adequately evaluate and respond to the individual needs of their struggling students. This practice creates stronger problem solvers and more flexible and responsive instructors who can adapt to the wide-ranging abilities of their students. Just as we ask teachers to be ever vigilant in assessing student understanding and adapting to meet student needs, so, too, must coaches be increasingly aware of their responsibility to identify and remediate gaps in teachers’ understanding and practice. Coaches should also focus on providing modeling and responsive teaching to strengthen the knowledge base supporting teachers’ instruction.

Rather than act as a resource or only scaffold implementation of a curriculum, coaches should build reciprocal and collaborative relationships incorporating problem-solving techniques

and constructive reflection targeted on improving practices. Similar to the instruction of classroom students, coaches should provide a safe environment for practice and application of new information and guide teachers through cognitive shifts, promoting deeper understanding of the evidence-based principles behind classroom instructional practices. Also similar to the way teachers approach classroom instruction, coaches should shoulder the responsibility of responding to the needs of their “students,” the teachers, and work together with teachers to identify areas of growth as well as those that may require improvement. By so doing, coaches can foster autonomy, confidence, and independence within the teachers they support as they learn to self-monitor, reflect, evaluate and improve their own practice.

Future Research

To move forward in the field of coaching research we must address three questions: a) what does success look like?; b) how can we measure it?; and c) how do we attain it?

Future studies would benefit from larger sample sizes, with equally distributed participants per condition, and should serve to answer one of the following questions: a) Are we accurately measuring the effects of coaching on teacher development?; b) Are we clearly identifying which coaching practices make a measureable difference in teacher and student achievement?; c) Is one type, style, or amount of coaching more effective at improving teacher practice over another?; and d) Are other types of professional development more effective than coaching at improving student achievement outcomes or instruction?

Additional studies should serve to fill the gaps in the current literature and explore the types of professional development that are most effective at increasing teacher content and pedagogical knowledge. Further research is needed to create and validate tools to adequately

capture growth in teacher knowledge and development in practice. Well-developed measures could in turn be used to help identify the components of coaching that are impactful and provide a body of evidence-based coaching practices. This information could then be paired with the random assignment of teachers to different coaching conditions to determine the most efficient manner to support teachers in applying new knowledge and creating observable changes in instructional practice.

Studies of varying amounts of coaching support – examining frequency and duration of visits – and whether or not technology might play a role in increasing the effectiveness of coaches would add to the body of literature currently reviewing the practice. Follow-up studies that include maintenance data may also be necessary to explore whether any observable changes in practice are in place after coaches are no longer readily available. A mixed methods approach to the study of coaching may better delineate which coaching practices hold promise of measurable impact and which are only supported by social validity standards.

Limitations

Findings from this exploratory analysis should be interpreted and considered in light of limitations common to research studies conducted in traditional school settings. Generalizations should be applied with caution due to the restrictions in sample size, duration and frequency of coaching support, measurement and differences in participant-coach relationships.

The nature and need of individual institutions influenced both the number and randomization of participants. As is often the case with campus-level recruiting, principals and school demographics dictated the number of educators able to participate in the study. Three of included schools in this study only had two teachers of general education first-grade classrooms

available for participation in the study. As teachers were randomized by campus to control (n=8), coached (n=6) and un-coached (n=7) conditions, this fact left a disproportionate amount of participants in un-coached or control conditions. One challenge associated with randomizing participants in studies with small sample sizes is the likelihood that groups will be not be equivalent. Evidence of this in the present study was the significant pre-test differences of participants on the Teacher Knowledge Survey. Both the small sample size and the unbalanced number of participants per condition also decreased the power required for robust statistical analysis and increased the likelihood of Type II error.

A related restriction is the limited number of contacts and short duration of coaching visits. The pressing nature of statewide assessments limited the amount of time spent in and frequency of visits coaches were able to have with teachers in their classrooms. Coaching was consistently provided on a bi-weekly basis during the seventeen-week intervention as projected, yet the total number of visits ranged from seven to twelve visits per teacher, with one teacher requiring additional support, resulting in 14 visits. As the intervention study closed one week prior to the start of statewide assessments, no follow-up visits were conducted nor maintenance data obtained.

Another limitation common to mixed methods designs is the lack of standardly accepted measures, validated and reliable, to measure teacher knowledge and quantify observable behaviors in teaching and coaching practices. As no industry-standard measure exists, researcher-made measures were created and used to measure teacher knowledge, implementation fidelity and coaching fidelity. As is the case with many researcher-made measures, no available data exists on the reliability or validity of the teacher knowledge survey.

A limitation unique to the examination of the effects of coaching is the individualized nature of coaching itself. Whereas a continual effort was made to standardize coaching practices and procedures, coaches were also instructed to co-select improvement targets with their assigned teachers. This type of individualized, responsive coaching may have led to differences in the amount (frequency of visits) and type of coaching (modeling, observing, co-planning) received by each educator, making the coaching procedures themselves less generalizable and consistent across participants. Allowable inconsistencies in the coaching protocol served to strengthen coaching relationships as coaches individualized the amount and type of support required for each participant. Resistant participants, only one of which was reported, also provided barriers to coaching contact and instruction, limiting the effectiveness and ability of the coach to support educational practice.

Summary

The purpose of this study was to examine the effects of sustained coaching support on teacher knowledge, implementation fidelity and first-grade student reading achievement. Coached teachers did not have significantly higher gains on a measure of teacher knowledge but did show significantly higher fidelity ratings than educators who only received professional development training. While students who participated in the Tier II reading intervention did show significant gains on measures of word reading and comprehension over students who did not receive the intervention, students of coached teachers did not perform significantly higher than those of un-coached teachers. Considering the high social validity coaching studies often report, this study may cause further reflection on the empirical evidence for coaching support.

Appendix A

Reading Rules Coaching Fidelity Form 2012-2013

Coaching Skill	Rating				
	Not Present	Partially Present	Present	N/A	Comments/Notes
1. Coach observed entire lesson. (all information that was presented)					
2. Coach collected pertinent data on students and/or teachers.					
3. Coach provided in-lesson support/scaffolding/modeling as needed during the lesson.					
4. Coach used data to drive/inform post-observation feedback.					
5. Coach provided positive feedback/praise citing specific examples or strengths.					
6. Coach indicated areas to be strengthened/parts of the lesson that need improvement.					
7. Coach listened to teacher questions/concerns.					
8. Coach addressed questions/concerns and provided resources as necessary.					
9. Coach and teacher co-selected target/goal for improvement.					
10. Coach suggested action steps to meet goal/target.					

Appendix B

Reading Rules Fidelity and Quality of Implementation 2012-2013

WHOLE GROUP COMPREHENSION ANCHOR LESSON

		Exp Y/N	Imp Y/N	Oral Presentation Adheres to Manual			Implements Instructional Activities as Described			Provides Appropriate Scaffolding & Feedback			Accuracy of Student Responding			Notes
				1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	
10-15 minutes																*If hand signal is supposed to be introduced it will be included in the introduction.
	Introduction	Y														
Start Time	Explicit Teach/ Review	Y														
End Time	Model & Guided Practice	Y														
	Independent Practice	Y														
Total Minutes	Closing & Review	Y														
	Pacing/Use of Instructional Time			1 Dragging or Rushed, A lot of Down Time			2 Adequate Pacing with some Down Time			3 Very Good Pacing with Little or No Down Time						
Unit	Teacher Preparation & Organization			1 Appeared Unprepared, Very Disorganized			2 Adequately Prepared and Organized, but With Several Lapses			3 Appeared Prepared, Well Organized with Few or No Lapses						
Lesson	Students Engaged & On-Task			1 Most students off-task			2 Some students on-task, Some students off-task			3 Most students on-task						

How challenging were the activities for most students in the group?	<input type="checkbox"/> Very Hard <input type="checkbox"/> About Right <input type="checkbox"/> Too Easy	Comments:
---	---	------------------

WHOLE GROUP COMPREHENSION READ ALOUD

		Exp Y/N	Imp Y/N	Oral Presentation Adheres to Manual			Implements Instructional Activities as Described			Provides Appropriate Scaffolding & Feedback			Accuracy of Student Responding			Notes
				1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	
10-15 minutes																
Start Time _____	<u>Before Reading:</u> Introduce/Review Review Skill or Story & States Guiding Question	Y														
End Time _____	<u>During Reading:</u> Covers in-text labels	Y														
Total Minutes _____	<u>After Reading:</u> Revisits Guiding Question	Y														
Unit _____	Pacing/Use of Instructional Time	1 Dragging or Rushed, A lot of Down Time			2 Adequate Pacing with some Down Time			3 Very Good Pacing with Little or No Down Time								
Lesson _____	Teacher Preparation & Organization	1 Appeared Unprepared, Very Disorganized			2 Adequately Prepared and Organized, but With Several Lapses			3 Appeared Prepared, Well Organized with Few or No Lapses								
	Students Engaged & On-Task	1 Most students off-task			2 Some students on-task, Some students off-task			3 Most students on-task								
How challenging were the activities for most students?		<input type="checkbox"/> Very Hard <input type="checkbox"/> About Right <input type="checkbox"/> Too Easy		Comments:												

SMALL GROUP COMPREHENSION

		Exp Y/N	Imp Y/N	Oral Presentation Adheres to Manual			Implements Instructional Activities as Described			Provides Appropriate Scaffolding & Feedback			Accuracy of Student Responding			Notes
				1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	
10 min.																
	Introduction															
Start Time	Explicit Teach/ Review															
_____	Model & Guided Practice															
End Time	Independent Practice															
_____	Closing & Review															
Total Minutes																
_____	Pacing/Use of Instructional Time			1 Dragging or Rushed, A lot of Down Time			2 Adequate Pacing with some Down Time			3 Very Good Pacing with Little or No Down Time						
Unit	Teacher Preparation & Organization			1 Appeared Unprepared, Very Disorganized			2 Adequately Prepared and Organized, but With Several Lapses			3 Appeared Prepared, Well Organized with Few or No Lapses						
Lesson	Students Engaged & On-Task			1 Most students off-task			2 Some students on-task, Some students off-task			3 Most students on-task						

How challenging were the activities for most students in		__Very Hard __About Right		Comments:												

the group?	__Too Easy	
------------	------------	--

		Exp Y/N	Imp Y/N	Oral Presentation Adheres to Manual			Implements Instructional Activities as Described			Provides Appropriate Scaffolding & Feedback			Accuracy of Student Responding			Notes
				1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	
10 min.																
Start Time	Phonemic Awareness															
_____	Letter Sound Instruction															
End Time	Sound Analysis															
_____	Decoding															
Total	HFW Instruction Recognition															
Minutes																

Unit	Pacing/Use of Instructional Time	1 Dragging or Rushed, A lot of Down Time	2 Adequate Pacing with some Down Time	3 Very Good Pacing with Little or No Down Time
Lesson	Teacher Preparation & Organization	1 Appeared Unprepared, Very Disorganized	2 Adequately Prepared and Organized, but With Several Lapses	3 Appeared Prepared, Well Organized with Few or No Lapses
	Students Engaged & On-Task	1 Most students off-task	2 Some students on-task, Some students off-task	3 Most students on-task

WORD STUDY

How challenging were the activities for most students in the group?	<input type="checkbox"/> Very Hard <input type="checkbox"/> About Right <input type="checkbox"/> Too Easy	Comments:
---	---	------------------

TEXT READING

	Word Study Activities	Exp Y/N	Imp Y/N	Oral Presentation Adheres to Manual			Implements Instructional Activities as Described			Provides Appropriate Scaffolding & Feedback			Accuracy of Student Responding			Notes
				1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	1 (Low)	2 (Med)	3 (High)	
10 min.																* Sentence reading is only for units 1-16. GQ = Guiding Question
	Sentence Rdg.															
Start Time	Before Reading: Introduction, states GQ															
	During Reading: covers in-text labels, restates GQ															
End Time	After Reading: Praise point, teach point, discusses GQ															
Total Minutes	Day 2 Fluency: Teaches/Reviews Strategy															
	Pacing/Use of Instructional Time	1 Dragging or Rushed, A lot of Down Time			2 Adequate Pacing with Some Down Time			3 Very Good Pacing with Little or No Down Time								
Unit	Teacher Preparation & Organization	1 Appeared Unprepared, Very Disorganized			2 Adequately Prepared and Organized, but With Several Lapses			3 Appeared Prepared, Well Organized with Few or No Lapses								
Lesson	Students Engaged & On-Task	1 Most students off-task			2 Some students on-task, Some students off-task			3 Most students on-task								
How challenging were the activities for most students in the group?		<input type="checkbox"/> Very Hard <input type="checkbox"/> About Right <input type="checkbox"/> Too Easy			Comments:											

GLOBAL QUALITY OF INSTRUCTION & BEHAVIOR MANAGEMENT				
Overall Quality of Instruction	Compared to other reading intervention lessons you observed, this lesson was:	Below Average 1	About Average 2	Above Average 3
Overall Quality of Behavior Management	<p>Compared to other classrooms you have observed, the quality of behavior management in this lesson was:</p> <p>(NOTE: A group may have one or more students who have serious difficulties with self-control/attention; if teacher re-directed these students and appeared to have generally effective behavior management routines in place, the rating could still be 3.)</p>	Low Quality 1	Moderate Quality 2	High Quality 3

LESSON TIMING : Were lesson components implemented within the prescribed times?			
Large Group Comprehension (15 min)	Less than 10 min <u>or</u> more than 20 min 1	10-20 min (10, 11, 19 or 20 minutes) 2	12-18 min 3
Small Group Comprehension (10 min)	Less than 5 min <u>or</u> more than 15 min 1	5-15 min (5, 6, 14 or 15 minutes) 2	7-13 min 3
Word Study (10 min)	Less than 5 min <u>or</u> more than 15 min 1	5-15 min (5, 6, 14, or 15 minutes) 2	7-13 min 3
Text Reading (10 min)	Less than 5 min <u>or</u> more than 15 min 1	5-15 min (5, 6, 14 or 15 minutes) 2	7-13 min 3

IMPLICATIONS FOR THE PROGRAM

Did the **teacher** have any specific problems that might be alleviated or reduced through changes to the program or training? This includes problems with: Implementing the **script/oral delivery** of the lesson (keeping place, had to read because there was too much verbiage, etc.); Implementing the **instructional routines** (using manipulatives, finding what was needed, figuring out what to do next, etc.); **Finishing the lesson component in the prescribed time** (too many activities in the component, activities were too complex or difficult to implement, or too hard for the students); **Meeting the needs of students at different** levels in the same group. **Did the students have any particular problems with any activities?**

LESSON COMPONENT	COMMENTS: Describe problem and (if you can) suggest potential changes to program or training. (Continue on back of form or add another sheet if needed)
Large Group Comp.	
Small Group Comp.	
Word Study	
Text Reading	

Appendix C

Reading Rules! Teacher Knowledge Survey 2012-2013

Please check the box that reflects the best instructional reading practices for at-risk readers.

	<i>Strongly Disagree</i>	<i>Disagree</i>	<i>Neither agree or disagree</i>	<i>Agree</i>	<i>Strongly Agree</i>
1. One-on-one instruction is always better for instructing struggling readers than small-group instruction.					
2. It is not necessary for a child to know how to decode prior to teaching comprehension skills.					
3. If a child cannot retell the important parts of a story the best approach is to have them re-read the story.					
4. To best support comprehension, a story should be read without interruptions.					
5. When a student makes an error, the teacher should immediately tell him/her the correct word.					
6. To make sure all students are able to participate, is it best to use 'round robin' reading, and have each student read in turn.					
7. When a beginning reader stops at an unknown word the teacher should have other students provide the word for the reader.					
8. Asking questions throughout a read-aloud activity is more effective at building comprehension than waiting to ask questions at the end of the story					
9. Reading comprehension should almost always be taught in a "whole class" format					

since every child needs to master the same comprehension skills.					
10. It is important that students use context (i.e., “meaning cues”) to help them identify unknown words in a story.					

Imagine that you’ve been asked to serve as a mentor to a first-year teacher in your school. Please list a few of the following approaches you would encourage him/her to use during reading instruction.

11. List the three most important things you would advise the new teacher to do when teaching reading comprehension.

- a.
- b.
- c.

12. List the three most important things you would advise the new teacher to do when teaching decoding/word reading skills.

- a.
- b.
- c.

13. List three effective strategies the new teacher can use to promote reading fluency.

- a.
- b.
- c.

14. What should the teacher tell a student to do when the student comes to a decodable word they do not know?

15. What should the teacher do when a student guesses a word?

16. What should the teacher do when a student inserts a word that is not in the text?

17. Mrs. Leeth has been doing an activity in which she says a multisyllabic word aloud and then asks the first graders in her intervention group to segment it chorally into syllables. After several lessons she realizes that one student is struggling to segment appropriately. When Mrs. Leeth says “multiply” he says, /m/ /ul/ /ti/ /pli/. What can Mrs. Leeth do to help this student?

18. The ability to segment a spoken word into its phonemes and blend phonemes together is necessary in order to

- a. sound out and spell words
- b. recognize sight words instantly
- c. comprehend text
- d. read text fluently
- e. write a story

19. Which of the following is the most useful for teaching a student to read the word *some*?

- a. Teach the silent *e* rule
- b. Teach the short *o* sound.
- c. Teach the student to sound out the word and blend the sounds from left to right.
- d. Teach the student to look closely at the word so that they can recognize it at sight and provide practice reading it quickly.
- e. Teach the short *u* sound.

20. You are teaching a new letter-sound. Which of the following is the process you should use?
- a. Ask the students the sound of the letter; tell them the correct sound; have all students say the new sound together with the teacher; finally mix the new sound in with other sounds they have learned and have the students practice all the sounds, both together and individually.
 - b. Wait until the student is reading and see if they miss a word with the letter in it before you teach it.
 - c. Ask the students the sound of the letter; have each child give the sound by themselves; finally mix the new sound in with other sounds they have learned and have the students practice all the sounds.
 - d. Tell the students the sound of the letter; have all students say the new sound together with the teacher and as a group without the teacher; have each child give the sound by themselves; finally mix the new sound in with other sounds they have learned and have the students practice all the sounds, both together and individually.
 - e. Ask each student the sound of the new letter and correct them right away if they answer incorrectly.

21. As students read a decodable text during her intervention, Ms. Harris notices that all of the students add a schwa sound to the end of some letter sounds, making the decoding of blends difficult. For example, one student tries to read the word 'plan' but says /pu/ /la/ /nu/. What might she do to correct this problem?

22. When a student makes a mistake during text reading
- a. The teacher should let the student keep reading and not interrupt him/her.
 - b. The teacher should stop the student and correct the error; then the student should go back to the beginning of the sentence and start again.
 - c. The teacher should have the student say the word correctly several times, then have the group repeat the word correctly before proceeding
 - d. The tutor should write down the word that was missed so it can be taught after students finish reading the story.
 - e. a. and d.

REFERENCES

- Ackland, R. (1991). A review of the peer coaching literature. *Journal of Staff Development, 12*, 22-27.
- Armbruster, B., Lehr, F., & Osborn, J. (2001). *Put reading first*. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement (CIERA).
- Armstrong, K., Cusumano, D. L., Todd, M., & Cohen, R. (2008). Literacy training for early childhood providers: Changes in knowledge, beliefs, and instructional practices. *Journal of Early Childhood Teacher Education, 29*(4), 297-308.
- Aud, S. & Hannes, G. (Eds.) (2011). The condition of education in 2011 in brief. (NCES 011—34). U.S. Department of Education, National Center for Educational Statistics. Washington, DC: U.S. Government Printing Office.
- Aud, S., Hussar, W., Kena, G., Bianco, K., Frohlich, L., Kemp, J., Tahan, K. (2011). The Condition of Education 2011 (NCES 2011-033). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Bean, R. M., Draper, J. A., Hall, V., Vandermolen, J., & Zigmond, N. (2010). Coaches and coaching in Reading First schools: A reality check. *The Elementary School Journal, 111*(1), 87-114.
- Biancarosa, G., Bryk, A. S., & Dexter, E. R. (2010). Assessing the value-added effects of

- literacy collaborative professional development on student learning. *The Elementary School Journal*, 111(1) 7-34.
- Birman, B. F., Boyle, A., Le Floch, K. C., Elledge, A., Holtzman, D., Song, M., ... & Yoon, K. S. (2009). *State and Local Implementation of the "No Child Left Behind Act." Volume VIII--Teacher Quality under "NCLB": Final Report*. Washington, DC: U.S. Department of Education; Office of Planning, Evaluation and Policy Development; Policy and Program Studies Service.
- Birman, B., Le Floch, K. C., Klekotka, A., Ludwig, M., Taylor, J., & Walters, K. (2007). *State and local implementation of the No Child Left Behind Act: Vol. 2. Teacher quality under NCLB: Interim report*.
- Bos, C., Mather, N., Dickson, S., Podhajski, B., & Chard, D. (2001). Perceptions and knowledge of preservice and inservice educators about early reading instruction. *Annals of Dyslexia*, 51, 97-120.
- Boyle, B., While, D. & Boyle, T. (2004). A longitudinal study of teacher change: what makes professional development effective? *The Curriculum Journal*, 15(1), 45-68.
- Bush, R. N. (1984). *Effective staff development. In making our schools more effective: Proceedings of three state conferences*. San Francisco: Far West Laboratories.
- Carlisle, J. F., Cortina, K. S., & Katz, L. A. (2011). First-grade teachers' response to

- three models of professional development in reading. *Reading and Writing Quarterly*, 27, 212-238.
- Carlisle, J., Kelcey, B., Rowan, B. & Phelps, G. (2011). Teachers Knowledge About Early Reading: Effects on Students' gains in reading achievement. *Journal of Educational Effectiveness*, 4(4), 289-321.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavior Sciences*, Psychology Press. Hillsdale, New Jersey: Psychology Press.
- Cohen, D. K. & Ball, D. L. (1999). *Instruction, capacity, and improvement*. (CPRE Research Report Series No. RR-43). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.
- Cornett, J. & Knight, J. (2009). Research on Coaching. In J. Knight (Ed.), *Coaching: Approaches and Perspectives* (192-216). Thousand Oaks, CA: Corwin Press, Inc.
- Cusumano, D. L., Armstrong, K., Cohen, R. & Todd, M. (2007). Indirect impact: How early childhood educator training and coaching impacted the acquisition of literacy skills in preschool students. *Journal of Early Childhood Teacher Education*, 27, 363-377.
- Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence*. Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.

- Darling-Hammond, L. (2008). Teacher learning that supports student learning. *Teaching for Intelligence*, 2, 91-100.
- Denton, C. A. (2012). Response to Intervention for Reading Difficulties in the Primary Grades Some Answers and Lingering Questions. *Journal of learning disabilities*, 45(3), 232-243.
- Domitrovich, C. E., Gest, S. D., Gill, S., Bierman, K. L., Welsh, J. A., & Jones, D. (2009). Fostering high-quality teaching with an enriched curriculum and professional development support: The Head Start REDI program. *American Educational Research Journal*, 46(2), 567-597.
- Edwards, J. L., & Green, K. E. (April, 1999). *Growth in coaching skills over a three-year period: Progress toward mastery*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.
- Elish-Piper, L. & L'Allier, S.K. (2011). Examining the relationship between literacy coaching and student reading gains in grades K-3. *The Elementary School Journal*, 112(1), 83-106.
- Foorman, B. R., & Moats, L. C. (2004). Conditions for sustaining research-based practices in early reading instruction. *Remedial and Special Education*, 25(1), 51-60.
- Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group

- instruction promote reading success in all children. *Learning Disabilities Research & Practice*, 16(4), 203-212.
- Fuchs, D., Fuchs, L. S., & Stecker, P. M. (2010). The "blurring" of special education in a new continuum of general education placements and services. *Exceptional Children*, 76(3), 301-323.
- Garmston, R., Linder, C. & Whitaker, J. (1993). Reflections on cognitive coaching. *Educational Leadership*, 51, 57-61.
- 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.
- Haycock, K. (1998). *Good teaching matters: How well-qualified teachers can close the gap*. ERIC Clearinghouse.
- Hsieh, W., Hemmeter, M. L., McCollum, J. A., & Otrosky, M. M. (2009). Using coaching to increase preschool teachers' use of emergency literacy strategies. *Early Childhood Research Quarterly*, 24, 229-247.
- International Reading Association. (2010). *Standards for reading professionals—Revised 2010*.
Newark, DE: Author.

International Reading Association. (2006). *Standards for middle and high school literacy coaches*. Newark, DE: Author.

International Reading Association, & National Council of Teachers of English. (2006). *Standards for middle and high school literacy coaches*. International Reading Association.

Ippolito, J. (2010). Three ways that literacy coaches balance responsive and directive relationships with teachers. *The Elementary School Journal*, 111(1), 164-190.

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development*. ASCD.

Knight, J. & Cornett, J. (n.d.). *Studying the impact of instructional coaching*. Lawrence, KS: University of Kansas Center for Research on Learning.

Knight, J. (2007). *Instructional coaching: A partnership approach to improving instruction*. Thousand Oaks, CA: Sage.

Knight, J. Ed. (2009). *Coaching: Approaches & perspectives*. Thousand Oaks, CA: Sage.

Kohler, F. W., Crilley, K. M., Shearer, D. D., & Good, G. (1997). Effects of peer coaching on teacher and student outcomes. *The Journal of Educational Research*, 90(4), 240-250.

Kretlow, A. G., & Bartholomew, C. C. (2010). Using coaching to improve the fidelity of

- evidence-based practices: A review of the studies. *Teacher Education and Special Education*, 33(4), 279-299.
- Kretlow, A. G., Wood, C. L., & Cooke, N. L. (2009). Using in-service and coaching to increase kindergarten teachers' accurate delivery of group instructional units. *Journal of Special Education*, 44, 234-246.
- Lane, Hudson, Leite, Kosanovich, Strout, Fenty & Wright (2009). Teacher knowledge about reading fluency and indicators of students' fluency growth in reading first schools. *Reading and Writing Quarterly*, 25(1), 57-86.
- Lockwood, J. R., McCombs, J. S., & Marsh, J. (2010). Linking reading coaches and student achievement: Evidence from Florida Middle Schools. *Educational Evaluation and Policy*, 32, 372-388.
- Lonigan, C. J., & Shanahan, T. (2008). Developing Early Literacy: Report of the National early literacy panel: A Scientific Synthesis of Early Literacy Development and Implications for Intervention Web site at [www. nifl. gov](http://www.nifl.gov). National Institute for Literacy.
- Lyon, G. R., & Weiser, B. (2009). Teacher knowledge, instructional expertise, and the development of reading proficiency. *Journal of learning disabilities*, 42(5), 475-480.
- Marsh, J. A., McCombs, J. S., & Martorell, F. (2010). How Instructional Coaches

- Support Data-Driven Decision Making Policy Implementation and Effects in Florida Middle Schools. *Educational Policy*, 24(6), 872-907.
- Mather, N., Bos, C., & Babur, N. (2001). Perceptions and Knowledge of Pre-service and In-service Teachers about early literacy. *Journal of Learning Disabilities*, 34(5), 472-482.
- Matsumura, L. C., Garnier, H. E., Correnti, R., Junker, B., & Bickel, D. D. (2010). Investigating the effectiveness of a comprehensive literacy coaching program in schools with high teacher mobility. *The Elementary School Journal*, 111(1), 35-62.
- Matsumara, L. C., Sartoris, M., Bickel, D. D., & Garnier, H. E. (2009). Leadership for literacy coaching: The principal's role in launching a new coaching program. *Educational Administration Quarterly*, 45, 655-693.
- McNamara, J. K., Scissons, M., & Gutknecht, N. (2011). A Longitudinal Study of Kindergarten Children At Risk for Reading Disabilities The Poor Really Are Getting Poorer. *Journal of learning disabilities*, 44(5), 421-430.
- Moats, L. (1994). The missing foundation in teacher education: Knowledge of the structure of spoken and written language. *Annals of Dyslexia*, 44(1), 81-102.
- Moats, L. (2009a). Knowledge foundations for teaching reading and spelling. *Reading and Writing*, 22(4), 379-399.

- Moats, L. (2009b). Still wanted: Teachers with knowledge of language. *Journal of Learning Disabilities*.
- Moats, L. C., & Foorman, B. R. (2003). Measuring teachers' content knowledge of language and reading. *Annals of Dyslexia*, 53(1), 23-45.
- National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.
- Neufield, B., & Roper, D. (2003). *Coaching: A strategy for developing instructional capacity—Promises and practicalities* (Prepared for The Aspen Institute Program on Education and The Annenberg Institute for School Reform). Cambridge, MA: Education Matters.
- Neuman, S. B. & Cunningham, L. (2009). The impact of professional development and coaching in early language and literacy instructional practices. *American Education Research Journal*, 46(2), 532-566.
- Neuman, S. B., & Wright, T. S. (2010). Promoting language and literacy development for early childhood educators: A mixed-methods study of coursework and coaching. *The Elementary School Journal*, 111(1), 63-86.

No Child Left Behind Act of 2001, 20 U.S.C. § 6319 (2008).

Phelps, G. & Schilling, S. (2004). Developing measures of content knowledge for teaching reading. *The Elementary School Journal*, 105 (1), 31-48.

Rabun, A., and West, J. (2004). *From Kindergarten Through Third Grade: Children's Beginning School Experiences* (NCES 2004–007), table A-9 and previously unpublished tabulation.

Reschly, A. L. (2010). Reading and school completion: Critical connections and Mathew effects. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, 26(1), 67-90.

Ross, J. A. (1992). Teacher efficacy and the effects of coaching on student achievement. *Canadian Journal of Education*, 17(1), 51-65.

Rupley, W. H. (2011). Research on teacher quality: Improving reading and writing instruction. *Reading & Writing Quarterly*, 27(3), 179-182.

Sailors, M. & Price, L. R. (2010). Professional development that supports the teaching of cognitive reading strategy instruction. *The Elementary School Journal*, 110(3), 301-322.

Shidler, L. (2008). The impact of time spent coaching for teacher efficacy on student achievement. *Early Childhood Education Journal*, 36, 453-460.

- Shidler, L. (2009). The impact of time spent coaching for teacher efficacy on student achievement. *Early Childhood Education Journal*, 36, 453-460.
- Showers, B. & Joyce, B. (1996). The evolution of peer coaching. *Educational Leadership*, 53(6), 12-17.
- Showers, B., Joyce, B., & Bennett, B. (1987). Synthesis of research on staff development: A framework for future study and a state-of-the-art analysis. *Educational Leadership*, 45, 77-87.
- Snow, C. E., Griffin, P., & Burns, M. S. the NAE Subcommittee on Teaching Reading. (2005). *Knowledge to support the teaching of reading: Preparing teachers for a changing world*.
- Snow, C., Griffin, P., & Burns, M. S. (Eds.). (2007). *Knowledge to support the teaching of reading: Preparing teachers for a changing world*. John Wiley & Sons.
- Swanson, H. L., & Hoskyn, M. (2001). Instructing adolescents with learning disabilities: A component and composite analysis. *Learning Disabilities Research & Practice*, 16(2), 109-119.
- Teemant, A., Wink, J., & Tyra, S. (2011). Effects of coaching on teacher use of sociocultural instructional practices. *Teaching and Teacher Education*, 27, 683-693.
- Torgesen, J. K. (2000). Individual differences in response to early interventions in

- Reading: The lingering problem of treatment resisters. *Learning Disabilities Research and Practice*, 15(1), 55-64.
- Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of school psychology*, 40(1), 7-26.
- Torgesen, J. K. (2004). Preventing early reading failure. *American Educator*, 28 (3) 6-9.
- Vanderburg, M., & Stephens, D. (2010). The impact of literacy coaches: What teachers value and how teachers change. *The Elementary School Journal*, 111(1), 141-163.
- Vaughn, S., Gersten, R., & Chard, D. J. (2000). The underlying message in LD intervention research: Findings from research syntheses. *Exceptional Children*.
- Walpole, S., McKenna, M. C., Uribe-Sarain, Lamitina, D. (2010). The relationships between coaching and instruction in the primary grades: Evidence from high-poverty schools. *The Elementary School Journal*, 111(1), 115-140.
- Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational researcher*, 37(8), 469-479.
- What Works Clearinghouse Procedures and Standards Handbook v 2.1 (2008). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, What Works Clearinghouse.

Woodcock, Richard W., K. S. McGrew, and N. Mather. *Woodcock-Johnson tests of achievement*. Itasca, IL: Riverside Publishing, 2001.

Zwart, R. C., Wubbels, T., Bergen, T. C., & Bolhuis, S. (2007). Experienced teacher learning within the context of reciprocal peer coaching. *Teachers and Teaching: theory and practice*, 13(2), 165-187.

Zwart, R. C., Wubbels, T., Bergen, T., & Bolhuis, S. (2009). Which characteristics of a reciprocal peer coaching context affect teacher learning as perceived by teachers and their students?. *Journal of Teacher Education*, 60(3), 243-257.

VITA

Christa Dawn Haring was born in Hot Springs, Arkansas. After graduating from Cheyenne East High School in Cheyenne, WY she completed a Bachelor of Science Degree in audiology and speech-language Pathology and post-baccalaureate teaching certificate in special education from Brigham Young University. She proceeded to teach students with learning disabilities and behavior disorders in Utah, Georgia and Ohio before completing her Master of Science Degree in communication disorders at Florida State University. After serving as a speech-language pathologist in Florida for two years she began her doctoral program at the University of Texas at Austin.

Permanent Email Address: christa.haring@gmail.com

This dissertation was typed by the author.