

Abstract

Title: INSTRUCTIONAL BEHAVIORS AND STUDENT
READING OUTCOMES IN A SCRIPTED TIER 2
INTERVENTION FOR FOURTH GRADE
STRUGGLING READERS

Elizabeth Ann Montanaro Ph.D., 2012

Directed By: Professor Rebecca D. Silverman, Department of
Counseling, Higher Education and Special Education

This study examined instructors' spontaneous responses to student errors in a scripted Tier 2 reading intervention and the relationship between tutor responses and student comprehension outcomes. A sequential exploratory mixed-methods design was used to identify the types of behaviors tutors exhibited in response to student errors, using transcripts of lessons. Tutors used four types of off-script behaviors when reacting to students: (a) scaffolding, (b) telling, (c) unclear feedback, and (d) erroneous feedback. Differences in how tutors implemented the standard protocol were analyzed qualitatively and described using frequency counts. Tutors exhibited differences in the frequencies of each of the behaviors, and differed in how closely they adhered to scripted lesson. Although tutors overall exhibited 76.3% fidelity of implementation, certain components

of the lessons were frequently omitted –modeling of the strategy, describing the purpose of the strategy, and providing opportunities for practice. These omissions may have influenced overall responsiveness for students receiving intervention. To determine how tutor differences might influence student outcomes, the frequency counts of the four spontaneous tutor behaviors were entered into regression equations to predict posttest scores on three measures of reading comprehension –Maze, Gates MacGinitie, and ASKIT. Findings indicate that scaffolding was related to student growth on one curriculum based measure of reading comprehension. The other three behaviors –telling, unclear feedback, and erroneous feedback –were not significantly related to student outcomes. Limitations, in light of these findings are considered. Implications for planning intervention studies and tutor training, as well as future directions for research, are discussed.

INSTRUCTIONAL BEHAVIORS AND STUDENT READING OUTCOMES
IN A SCRIPTED TIER 2 INTERVENTION FOR FOURTH GRADE
STRUGGLING READERS

by

Elizabeth Ann Montanaro

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Professor Rebecca D. Silverman, Chair
Dr. Lisa Pericola Case
Professor Joan Lieber
Professor Kristen D. Ritchey
Professor Linda Valli

© Copyright by

Elizabeth A. Montanaro

2012

Acknowledgements

I would like to express my appreciation for the members of my committee whose feedback helped develop my research and writing. First, warmest thanks to Dr. Deborah Speece, Dr. Rebecca Silverman, Dr. Kristen Ritchey, and Dr. Lisa Pericola Case for their willingness to share their data and allow me to become a member of their team. It has been a privilege to work alongside such accomplished scholars and I am so grateful for the experience. Thank you to Dr. Joan Lieber whose critical feedback and keen insights have improved my work at every turn. Thank you to Dr. Linda Valli who gave me the tools I needed to complete a mixed methods study and asked me tough questions that helped steer my thinking. I especially wish to express my deepest appreciation to my advisor, Dr. Rebecca Silverman, who has read draft after draft of this manuscript and encouraged me to keep at it.

A special thanks to Amanda Stefanski, who served as a “critical friend,” assisted with coding, and provided her proofreading expertise. My work has benefitted from your watchful eye, insightful questions, and consistent support. I am grateful to call you a colleague and friend.

Finally, I must express my deepest gratitude to my family who has cheered me on throughout this process. Thank you to my mother who has spent hours and hours babysitting so I could work. To Jack, mommy is finally done with her “dissertation”! And of course, I could not have made it work without my husband who has supported me at every turn. As a writer, he was keenly aware of the painstaking process of drafting, writing, revising, and revising some more. Domenico, I am forever grateful for your unfailing understanding and encouragement. We make quite a team.

TABLE OF CONTENTS

	<i>Page</i>
Acknowledgements.....	ii
List Of Tables.....	iv
List of Appendices.....	v
CHAPTER I: INTRODUCTION TO THE PROBLEM.....	1
Response to Intervention.....	2
Statement of the Problem.....	5
Significance and Rationale of Current Study.....	16
Research Questions.....	17
CHAPTER II: REVIEW OF THE LITERATURE.....	19
Theoretical Background.....	19
Review of Empirical Studies.....	31
Content Review.....	34
Methodological Review.....	46
Methodological Considerations for Future Research.....	66
Synthesis and Directions for Future Research.....	67
CHAPTER III: METHOD.....	70
Purpose of the Study.....	70
Context.....	71
Current Study.....	74
Participants.....	76
Design.....	81
Qualitative Analysis.....	85
Quantitative Analysis.....	91
Summary.....	103
CHAPTER IV: RESULTS.....	105
Research Question 1.....	107
Research Question 2.....	125
Research Question 3.....	143
Research Question 4.....	149
Summary of Findings.....	157
CHAPTER V: DISCUSSION.....	165
Types of Spontaneous Instructional Behaviors.....	166
Fidelity of Implementation.....	179
Other Considerations.....	182
Implications.....	186
Limitations.....	188
Summary.....	190
Appendices.....	191
References.....	231

List of Tables

<i>Number</i>		<i>Page</i>
Table 1.	Tutor Demographics.....	79
Table 2.	Student Demographics.....	81
Table 3.	Behaviors with Relationships with Student Achievement.....	86
Table 4.	Predictors for GMRT Comprehension and Maze.....	101
Table 5.	Predictors for ASKIT.....	103
Table 6.	Frequency of Instructional Behaviors.....	108
Table 7.	Types of Scaffolding.....	109
Table 8.	Mean Frequency of Behaviors Across Lessons.....	127
Table 9.	Tutor Experience and Scaffolding Repertoire.....	128
Table 10.	Fidelity of Implementation Summary.....	144
Table 11.	Descriptive Statistics.....	152
Table 12.	Correlation Matrix.....	153
Table 13.	Summary of VIF Values.....	154
Table 14.	Regression Analysis for Predicting GMRT Comprehension Scores....	159
Table 15.	Regression Analysis for Predicting Maze Scores.....	161
Table 16.	Regression Analysis for Predicting ASKIT Scores.....	163

List of Appendices

<i>Appendix</i>		<i>Page</i>
A	Summary of Literature Included in Review.....	191
B	Example Lesson Plan.....	199
C	Intervention Scope and Sequence.....	205
D	Code Book.....	213
E	ASKIT Assessment with Scoring Examples.....	216
F	Fidelity of Implementation Data Sheets	230

Chapter I: Introduction to the Problem

Reading comprehension is an essential skill for success. Poor reading skills have been linked to poverty (Barton & Jenkins, 1995), higher rates of incarceration (Newman, Lewis & Beverstock, 1994; Svensson, Lundberg & Jacobson, 2003), as well as poor employment outcomes (Sum, Kirsch, & Taggart, 2002). One important influence on children's literacy development is the quality of instruction they receive (Morrison, Bachman, & Connor, 2005; Snow, 2002). Comprehension instruction has been proven to be an important influence on children's reading development (Aarnoutse, VanLeeuwe, Voeten, & Oud, 2001; Dickinson & DeTemple, 1998; Snow, 2002). In fact, the RAND Study Group (Snow, 2002) cited good instruction as the "most powerful means of developing proficient comprehenders and preventing reading comprehension problems" (p. xvii), which begs the question: what is good reading comprehension instruction?

The literature on effective comprehension instruction indicates that the explicit teaching of strategies is particularly beneficial for struggling readers (Connor, Morrison, & Petrella, 2004; Pressley & Wharton-McDonald, 1997; Snow, 2002). Research suggests that direct explanations paired with modeling and think alouds are effective means of teaching reading comprehension (Book, Duffy, Roehler, Meloth, & Vavrus, 1985; Duffy et al., 1986). Similarly, a synthesis of research literature on comprehension instruction for

struggling readers indicated that explicit modeling by the teacher and extensive teacher feedback are essential elements of quality comprehension instruction (Gersten, Fuchs, Williams, & Baker, 2001). Additionally, researchers have identified scaffolded practice of comprehension strategies to be an effective means of developing children's reading comprehension, in addition to direct explanations and modeling (Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994; Pressley et al. 1992; Rodgers, 1999).

Tutoring has been identified as a successful approach of instructional delivery for struggling readers. Multiple studies have explored the use of one-on-one or small group tutoring to improve the reading skills of struggling readers (Elbaum, Vaughn, Hughes & Moody, 2000; Slavin, Lake, Davis, & Madden, 2011). Given the recent emphasis on the use of Response to Intervention (RTI) models as a means of providing early intervention to struggling students, there has been renewed research interest in identifying critical aspects of successful tutoring (Slavin et al., 2011).

Response to Intervention

Response to Intervention (RTI) has been heralded as a promising alternative to the traditional discrepancy method in identifying learning disabilities (LD), while providing early intervention for all struggling students. School systems across the nation are rapidly adopting RTI models of intervention and identification (Reynolds & Shaywitz, 2009).

Thirty-seven states are currently implementing or in the process of developing RTI models of instruction and assessment (Berkeley, Bender, Peaster, & Saunders, 2009). A key assumption of RTI is the majority of students exposed to high-quality, research-based instruction will respond positively. Non-responsiveness to sound instruction and lack of progress despite increasingly intensive interventions may indicate the presence of LD. Although some researchers are in support of using RTI to identify and diagnose students with LD, many have voiced concerns about the validity and utility of this approach (Berkeley et al., 2009).

Gerber (2005) cautioned that if RTI is used as a method of identification, reliability could be compromised due to variance inherent to the practitioner delivering the intervention. Various characteristics of teachers have been linked to student achievement including certification status (Clotfelter, Ladd, & Vigdor, 2006; Heck, 2007; Rockoff, 2004); years of teaching experience (Darling-Hammond, 2000; Hanushek, Kain, O'Brien, & Rivkin, 2005; Rockoff, 2004); content knowledge (Darling Hammond, 2000; Goldhaber & Brewer, 2000); and other personal characteristics (Pianta et al., 2008; Sammons et al., 2007) In fact, Marzano (2001) determined that teacher skill accounted for increases in student performance on curriculum-based assessments by 20-45 percentage points. Although the literature supports the impact of teacher characteristics

on student outcomes, little research has considered the impact of instructor variability on student outcomes in small group interventions within an RTI model.

Interventions within an RTI framework often include detailed lesson plans or scripts for the instructor to follow in order to ensure treatment integrity. Variations in the fidelity of implementation must be considered in order to determine if a student's lack of progress was due to failure to implement the program as designed (O'Donnell, 2008). Fidelity of implementation, or treatment integrity, involves the *accuracy* and *consistency* with which intervention procedures are implemented (Wood, Umbreit, Liaupsin, & Gresham, 2007). The instructor's level of adherence to the treatment protocol has been shown to relate to student outcomes (Gresham, Gansle, Noell, Cohen, & Rosenblum, 1993; Kovalski, 2007), underscoring the necessity to consider other instructor-related behaviors that may impact responsiveness. Although the literature supports the impact of teacher characteristics on student outcomes, no research has considered the impact of instructor variability on student outcomes as a result of small group interventions delivered within an RTI model.

Even when provided with a tightly scripted intervention, variation remains in how the instructor responds to the student in the learning process. Although a question may have only one correct answer, the instructor can respond to both correct and incorrect

answers in a multitude of ways. Herein lies the variability that teachers lend to student outcomes even within highly controlled interventions. If decisions about student disability status are to be made as a result of tutoring outcomes, then the differential impact of spontaneous instructional behaviors on student outcomes must be examined, and accordingly accounted for, when defining and determining responsiveness.

Statement of the Problem

RTI and identification. In 1976, the U.S. Department of Education (USDOE) provided the first federal definition of learning disabilities, which included the criterion of a severe discrepancy between achievement and intellectual ability. Federal regulations did not specify what characterized a “severe discrepancy” or how to determine if such a discrepancy existed. In response, a formula was devised to substantiate the presence of a severe discrepancy, comparing measures of IQ to measures of academic achievement (Hammill, 1990). Subsequently, many states and districts adopted mathematical formulas to determine the presence of such a discrepancy, although this specific approach was not stated in federal regulations (Fuchs, Mock, Morgan, & Young, 2003; Hammill, 1990; Kavale, 2005). This approach was widely criticized, due to variations defining “severe,” measurement flaws associated with the use of formulas, as well as the rapid increase in the number of students, particularly minority students, identified as LD (Hollenbeck,

2007). Additionally, the discrepancy approach was dubbed a “Wait and Fail Approach” (Reschly, 2003) because it delayed treatment in earlier years, despite the presence of symptoms of a disability, until later school years (e.g., third grade and above) when the impact on school performance was more serious. Often these older struggling students are more difficult to motivate and more challenging to remediate (Boardman et al., 2008).

When the Individuals with Disabilities Act (IDEA) was reauthorized in 2004, the statute explicitly addressed the discrepancy issue, stating, “a local educational agency shall not be required to take into consideration whether the child has a severe discrepancy between achievement and intellectual ability” (IDEA, 2004). Further, it included language permitting the use of a process-based approach using a student’s response to scientific, research-based intervention as a means of LD identification. This process has been commonly referred to as Response to Intervention (RTI). RTI has multiple objectives, including providing all students with appropriate, research-based instruction; preventing and identifying learning problems within general education through a structure of multi-tiered assessment and intervention; and reducing the number of students requiring services outside general education (Johnson, Mellard, Fuchs, & McKnight, 2006). The RTI process includes ensuring quality classroom instruction (Tier 1), universal screening, frequent monitoring of student progress, and intervention services

(Tier 2) that vary in intensity and scope for students who are not progressing as anticipated. Tier 2 interventions are typically conceptualized as targeted instruction delivered within small groups. Those students who still do not respond adequately may be eligible for Tier 3 interventions or special education, as data collected during the RTI process may serve as one factor in determining if a student has a specific learning disability (Burns, Jacob, & Wagner, 2008; Fuchs et al., 2003).

Although RTI can be conceptualized in many ways, the three-tier model is most frequently described in the literature. The first tier of the RTI model establishes adequate classroom instruction is in place, incorporating “scientifically validated” instructional methods, thus assuring that student difficulties cannot be attributed to ineffective classroom instruction (Johnson et al., 2006). According to Fuchs and Fuchs (1998), the first tier substantiates that classroom instruction is generally effective enough to allow for individual decisions about student responsiveness to be made. For example, if most of the students in a class are not making satisfactory growth in comparison to local or national norms, an intervention at the classroom level is warranted to increase the effectiveness of classroom instruction being provided.

Once the classroom environment has been deemed “sufficiently nurturing” (Vaughn & Fuchs, 2003), assessments at the classroom level are used to determine if any

students are not adequately responding to instruction. Students who are not progressing as expected in the general education classroom are targeted for further intervention. If a student is unresponsive to increasingly intensive interventions, identification of a disability may be considered. However, the RTI process invites school personnel to consider the possibility that those students who are non-responsive may not be receiving adequate instruction, before assuming that the student has a deficit or disability (Harry & Klingner, 2006). If failure due to inadequate instruction can be ruled out, students may be identified with a LD and considered eligible for special education services. For example, in one proposed model of identification, students who are dually discrepant, that is, performing one standard deviation below their peers on curriculum based measures and whose rate of growth throughout the intervention measured by slope is also one standard deviation below the slopes of their peers despite high-quality interventions, may be identified as having difficulty severe enough to require special education (Fuchs & Fuchs, 1998; Speece & Case, 2001). Other researchers have identified nonresponders using the components of dual discrepancy approach individually –using final status based on posttest scores (Torgeson et al., 2001) or slope of growth (Vellutino et al., 1996). Others have proposed the use of benchmark or cut off scores (Good, Simmons, & Kame'enui, 2001), in which a score on a criterion-based measure is determined as the

standard. Students who score above the benchmark are considered responders, while those who do not achieve the predetermined score are considered non-responders.

Whichever method is used to determine lack of adequate response to instruction, students who meet that criterion are referred for more intensive interventions (Tier 2 or Tier 3) or be referred for special education eligibility.

RTI and reading. RTI can be used across academic areas as well as for behavioral issues (Johnson et al., 2006); however, the most substantial literature on RTI has described its utility in addressing early reading difficulties (Wanzek & Vaughn, 2007). Less is known about RTI in the later grades (Vaughn et al., 2008). It is estimated that between 22% and 42% of students identified as having reading disabilities encounter difficulties for the first time in third grade or later (Kavale & Resse, 1992; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992). This phenomenon of has been referred to or the *fourth-grade slump* (Chall & Jacobs, 2003), which coincides with the shift from *learning to read* to *reading to learn* as children begin to use their reading skills as a means of learning in the content areas (Chall, Jacobs, & Baldwin, 1990, p. 11). Chall posited that students at this age begin to struggle with reading comprehension because the concepts become more abstract and the vocabulary less familiar. Moreover, some students lack the specialized background knowledge, vocabulary, and familiarity with

expository texts required to comprehend grade level materials (Sanacore & Palumbo, 2009).

Although many researchers have cited the importance of developing interventions for students in later elementary grades in response to late emerging comprehension difficulties (Sanacore & Palumbo, 2009; Suhr, Hernandez, Grimes, & Warschauer, 2010), reading intervention research has largely focused on improving decoding skills for students in primary grades (Wanzek & Vaughn, 2007). Further research is needed to develop comprehension interventions for students in later elementary grades and identify effective tutoring behaviors, particularly as RTI becomes an increasingly popular approach to identify and remediate reading difficulties.

Approaches to RTI. There are two main methods of small group intervention delivery within RTI: (a) use of a standard treatment protocol and (b) the problem solving approach. Standard protocol is a uniform method of instruction where instructors follow a set lesson plan. The problem-solving model is a less standardized, more individualized approach in which an intervention is designed for a particular student, similar to traditional pre-referral intervention procedures (Fuchs, Fuchs, & Compton, 2004; McKenzie, 2009; Wanzek & Vaughn, 2007).

Standard protocol approach. Within standard protocol interventions the same research-validated instructional procedures are used for all students targeted for intervention, although the intervention's focus and level of instruction may vary based on student need (Wanzek & Vaughn, 2007). Standard protocol interventions typically include standardized, or scripted, procedures designed to guard against threats to the integrity of the intervention (Johnson et al., 2006). The accuracy and consistency with which intervention procedures are adhered is referred to as fidelity of implementation, or treatment integrity (Wood et al., 2007). Adherence to the instructional protocol is essential for maintaining the internal validity of the intervention. In fact, Gresham, MacMillan, Beebe-Frankenberger, and Bocian (2000) noted that it is virtually impossible to assess the effectiveness of a treatment without knowing how faithfully the treatment was applied. Since the degree of treatment fidelity is often related to treatment outcomes, data must be collected on how faithfully the intervention was implemented (Gresham et al., 2000; Kovalesski, 2007).

Problem-solving approach. While standard protocols require adherence to standardized procedures, the problem-solving approach calls for teams of school-based professionals to develop and monitor individualized instructional plans to address the academic or behavioral concerns of students who are struggling. Educators select and

adapt evidence-based instructional practices to meet the needs of individual students.

These plans are dynamic, and may change as a result of continuous progress monitoring data, which measures student responsiveness to the selected teaching methods.

Comparing the approaches. While nearly all of the research literature on RTI describes the use of a standardized approach (Fuchs et al., 2004; Wanzek & Vaughn, 2008), problem solving is currently in more common use in school settings (Fuchs et al., 2004). Despite its widespread popularity, D. Fuchs and his colleagues (2004) described several drawbacks to the problem-solving approach. One major concern is the striking similarity to traditional pre-referral interventions, which Flugum and Reschly (1994) documented are rarely carried out with quality, if at all. Additionally, the practitioners using the problem solving approach have had difficulties documenting and adhering to the levels of fidelity necessary to be regarded as a scientifically based approach to identification (Fuchs et al., 2004). Carney and Stiefel (2008) also noted that a greater professional expertise is required for implementing the problem solving approach, due to the reliance on decision-making and continuously adjusting instructional approaches. The use of a standardized intervention allows for greater quality control, as training and assessment of fidelity is less complicated (Fuchs et al., 2004). Ensuring quality control of

interventions will be necessary as schools begin to use responsiveness as an indicator of disability status.

Using RTI to identify reading disability. Vellutino et al. (1996) conducted the first major study to examine the utility of considering student responsiveness to a reading intervention as a diagnostic tool for reading disability (RD) identification. Since then, the efficacy of the standard protocol approach in early reading interventions has been well documented. Standard protocol interventions have been used in several studies to identify non-responders and remediate struggling students through extensive reading interventions (Mathes et al., 2005; Torgesen et al. 1999; Vellutino et al., 1996; Vellutino, Scanlon, Small, & Fanuele, 2006).

Issues with implementation. Regardless of the approach, it is important to consider instructional quality and the potential for variation. Standard protocol interventions in particular require strict adherence to the instructional protocol to ensure instructional quality and limit the impact of instructor variation. Many researchers studying reading interventions have expressed concerns over the strength of their findings due to variations in tutor or teacher training, skill, or expertise. For instance, Torgeson and his colleagues (1999) hypothesized that using paraprofessionals to alternate sessions with tutors, while economically beneficial, may ultimately have not been effective and

provided one possible explanation for the limited reading achievement of children in their study. Although the paraprofessionals were well-trained, they were not able to provide the type of practice activities needed by the struggling readers in the study (Torgeson et al., 1999). However this study did not directly compare the impact of using paraprofessionals versus tutors. And while Fuchs and Fuchs (2009) have emphasized that RTI depends on instruction from a highly skilled instructor with strong content knowledge and clinical expertise, it remains unclear how schools or researchers plan on accounting for "variable student achievement as a function of instructional effort" (Gerber, 2005, p. 517). Daly, Martens, Barnett, Witt and Olson (2007) have similarly emphasized the importance of the teacher's response to the student in the learning process, saying that effective teaching is responsive to student responses and their changes over time. In fact, Klingner and Edwards (2006) warn that given the vast differences in the level of knowledge, skills, temperament, and beliefs across teachers, it is simply not sensible to make the assumption that all teachers can implement interventions with sufficient quality to promote adequate opportunities for student learning. If decisions about student disability status are to be made based on intervention outcomes, the instructors must be highly trained and knowledgeable, particularly if the standard protocol allows for any decision-making.

Research on instructional variation. It is well documented that teachers differ in the way they implement curricula and adopt new instructional practices (Datnow & Castellano, 2000; Englert & Tarrant, 1995; Jenkins & Leicester, 1992; Klingner, Vaughn, Hughes & Arguelles, 1999). When implementing curricula, instructors tend to make adaptations depending on their understanding of content, beliefs about teaching and learning, perceived flexibility or lack of accountability, or perceptions about how the curricula is meeting students' needs (Datnow & Castellano, 2000). Although there is a small, but growing corpus of literature on the adaptations teachers make to curricula, little is known about which adaptations compromise student outcomes and which changes are advantageous.

Research on teacher effects. Changes in policy aiming to increase accountability and student achievement such as No Child Left Behind (2001) and Race to the Top (2009) have also spurred a resurgence of research examining the relationship between teacher variables and student achievement. Research has shown that some teachers are more effective than others. In fact, Kratochwill, Volpiansky, Clements, and Ball (2007) found that the effect of teacher variables on student achievement was second only to the influence of home-related factors. Teacher characteristics which have been associated with changes in student outcomes in reading and math include certification status

(Clotfelter, Ladd, & Vigdor, 2007; Heck, 2007; Rockoff, 2004); years of teaching experience (Darling-Hammond, 2000; Hanushek et al., 2005; Rockoff, 2004); and content knowledge as measured through major or number of courses completed in an area of study (Darling-Hammond, 2000; Goldhaber & Brewer, 2000). In addition to education, certification, or experience, researchers have found positive correlations between the personal characteristics of teachers and their students' achievement. For instance, commitment and resiliency (Sammons et al., 2007), and emotional support from teachers (Pianta et al., 2008) have also been linked to higher scores on measures of academic achievement. Although many researchers have demonstrated a relationship between student outcomes and teacher characteristics, few have considered the impact of variability in instructional practices on student outcomes in an RTI model.

Significance and Rationale of Current Study

School systems nationwide are rapidly adopting RTI models, yet many questions remain regarding the feasibility, practicality, and reliability of utilizing RTI as a method of LD identification in school-based settings (Kavale, Kauffman, Bachmeier, & LeFever, 2008; Mastropieri & Scruggs, 2005; McKenzie, 2009). Although the literature indicates that variations in student achievement can be associated with numerous teacher characteristics and behaviors, it is unclear if and how the influence of teacher-level

variables will be taken into account when labeling a student as responsive or unresponsive to instruction. A pressing area of research is to examine the instructional variation that exists across teachers even within standardized interventions. Gerber (2005) cited this gap in the literature, by pointing out that few RTI studies detail the variations in teachers' adjustments, decisions, and behaviors while implementing scripted interventions. Similarly, McGill-Franzen (2005) called for further study of teacher variation in adapting and implementing scientifically-based reading instruction, as the literature supports that teacher effects are large.

Research Questions

This study examined the types of instructional behaviors tutors exhibit while interacting with students during the implementation of a scripted reading intervention.

The research questions of interest are:

Research Question 1: What types of spontaneous instructional behaviors do tutors exhibit in reaction to student oral responses within a scripted reading comprehension intervention?

Research Question 2: How do tutors differ from each other in regard to the type, variety, and frequency of the identified spontaneous instructional behaviors?

Research Question 3: How do tutors differ from each other in terms of their fidelity of implementation of the scripted lesson plan?

Research Question 4: What is the relationship between the spontaneous instructional behaviors of tutors and student reading comprehension outcomes, controlling for student demographic variables, prior achievement, and fidelity of implementation?

Chapter II: Review of the Literature

The purpose of this chapter is to summarize the findings of studies exploring the role of instructional differences on student reading achievement. This chapter presents (a) relevant theory associated with student-instructor interactions, (b) two opposing theories of curriculum implementation and implications for RTI, (c) a brief overview of the existing research on implementation of interventions, (d) a content and methodological review of the existing literature on the relationship between student reading achievement and instructional behaviors, and finally, (e) implications for future research.

Theoretical Background

Social learning theory. Interactions between teacher and student are a key part of the process of teaching and learning (Palincsar, 1986; Vygotsky, 1978). Individuals acquire new behaviors within social contexts through observation and social interaction (Bandura, 1977; Palincsar, 1986), Social Learning Theory emphasizes the role of the teacher as "more knowledgeable other" (Mariage, Englert, & Garmon, 2000, p. 299) who should determine a student's developmental level using relevant measures, as well as gauge the student's current level of understanding and background knowledge through interactions with the student (Eagan, 2009). In turn, the teacher is able to present a task matched to a student's "zone of proximal development" (ZPD; Vygotsky, 1978, p. 33),

that is, the level at which a student cannot perform independently, but could accomplish in collaboration with more proficient peers or with adult support (Vygotsky, 1978).

Although commonly considered as espousing opposing perspectives, Bandura (1977) and Vygotsky (1978) both emphasized the role of socialization in learning.

Vygotsky is associated with the constructivist paradigm, in which the student constructs knowledge within the context of a collaborative relationship with the instructor (Powell & Kalina, 2009). Teachers and students work together, engaging in a dialogue in which both the teacher and the student learn (Palincsar, 1986). Bandura's work, on the other hand, is rooted in behaviorism. Here, socialization plays an equally important role, as Bandura cited social reinforcement and conformity as primary motives for engaging in imitative behaviors.

Bandura (1969) explained that most learning occurs through exposure to behavior models and verbal modeling cues. Observers of a modeled behavior often imitate the behavior they witnessed (Bandura & Barab, 1971). A verbal modeling cue is a stimulus, written or spoken, which provides information. For instance, one may acquire the ability to behave in a particular manner by reading a manual or following oral directions, without necessarily observing another person model the behavior (Bandura, 1969).

Behaviors can be reinforced via social reinforcement, which can be verbal (i.e., praise or

criticism) or non-verbal, (i.e., a smile, nod, or raised eyebrow; Bandura, 1969). Early studies of social reinforcement and learning demonstrated large, significant effects for verbal social reinforcement, particularly criticism (Allen, 1966; Allen, Spear, & Lucke, 1971; Stevenson & Cruse, 1961).

Although often falsely attributed to Vygotsky, the term "scaffolding" was actually introduced by Wood, Bruner, and Ross (1976, p. 90). Vygotsky's work served as the foundation for their conceptualization of scaffolding (Scrimsher & Tudge, 2003).

Scaffolding involves "the adult 'controlling' those elements that are beyond a learner's capacity" (Wood et al., 1976, p. 90), to gradually bridge the zone between what is known and unknown for the student. Scaffolding serves as a metaphor for temporary supports that can be adjusted as needed, and eventually removed (Palincsar, 1986), allowing students to perform tasks that would be unreachable without support. A continuum of strategies are involved in scaffolding (Pentimonti & Justice, 2010). When presenting a new skill or concept, the teacher initially provides strong supports, by demonstrating a task or thinking aloud (Wood et al., 1976). As students develop a skill, supports are gradually removed until minimal support is needed for the student to complete a task independently. Low levels of support may include verbal cues or prompts, with little adult assistance as the student nears mastery (Pentimonti & Justice, 2010). In addition to

modeling and scaffolding, Vygotsky (1987) defined a good teacher as one who “explains, informs, inquires, corrects, and forces the child himself to explain” throughout the instructional process (pp. 215-216). New concepts are co-constructed through social interaction, modeling, observation, and feedback. While social learning theory serves as the foundation of research on teacher-student interactions during instruction, the theoretical underpinnings concerning how instructors implement interventions must be considered within existing theories of curriculum implementation.

Theories of curriculum implementation. Implementation of curriculum can be viewed through two main theoretical frameworks –fidelity or adaptation (Cho, 1998). These are two polarized perspectives with diverging basic assumptions. Based in behaviorism, the paradigm of fidelity considers the extent to which the intervention is implemented as it was intended by the intervention developers (Cho, 1998). In contrast, the adaptation approach is associated with the theory of bounded rationality, which acknowledges factors exist beyond an individual's intellectual capacity—in this case, the intellectual capacity of the intervention designer—which impact how a program or its components are implemented. Adaptation advocates the selection and modification of program components by implementers to respond to variations in context and meet immediate needs (Emshoff et al., 1987).

Fidelity. The majority of the literature on RTI utilizes a fidelity-based approach using standardized tutoring protocols (Gersten et al., 2009). The creators of the intervention are considered experts, and teachers are expected to be "faithful, rational adopters" (Emshoff et al., 1987, p.301). A core assumption is that the curriculum was empirically validated as designed and any changes may result in less effective outcomes. Variation in fidelity of implementation influences the level of effectiveness (O'Donnell, 2008). To establish internal validity of an intervention, data concerning fidelity of implementation is essential. In fact, Gresham et al. (2000) explained it is virtually impossible to assess the effectiveness of a treatment without knowing how faithfully the treatment was applied. Since the degree of treatment fidelity is often related to treatment outcomes, data must be collected on how faithfully the intervention was implemented (Gresham, et al., 2000; Kovalski, 2007).

Adaptation. From an adaptation perspective, variation in local implementation is considered inevitable, even desirable. Adaptations are not seen as threats to the program's intentions, and are instead viewed as assets (Penuel & Means, 2004). Some researchers refer to this as the mutual adaptation model or co-construction perspective (O'Donnell, 2008). Problem solving models of RTI use an adaptation approach where interventions are continually tweaked based on student responsiveness (Carney & Stiefel, 2008).

Researchers have discovered that even well-defined curricular policies are implemented with substantial variation across settings, and are continually modified in the classroom, as teachers work to meet the needs of their students and the demands of school administration (Datnow & Castellano, 2000). Elmore and Sykes (1992) explained this variation might be due to perceived flexibility or lack of accountability in adhering to the approved curriculum or adopting mandated instructional methods. Datnow and Castellano (2000) discovered even teachers who strongly supported district-adopted reading curricula made adaptations during implementation. For instance, teachers admitted they modified the amount of time spent on individual components and omitted some activities entirely. For example, in a study of a scripted reading intervention program, Success for All (SFA), Klingner Cramer and Harry (2006) found that teachers who were rated as “effective” based on observations did not follow the SFA manual closely, often omitting parts of the program they felt were “time wasters”, adapting the amount of time spent on certain aspects of instruction, or modifying instructional procedures. In addition to preplanned adjustments, these skilled teachers were also able to make spontaneous adjustments in response to student understandings (Klingner et al., 2006). Although student outcomes were not directly measured in this study, the authors viewed these adaptations favorably, rather than threats to the efficacy of the intervention.

It is unclear at what point individual teachers' adaptations compromise the integrity and intent of the program (Datnow & Castellano, 2000). McGill-Franzen (2005) suggested researchers look more closely at the appropriateness of teacher adaptations and called for further study of the outcomes associated with adaptations made by teachers.

Background

Program implementation and student outcomes. The literature examining the relationship between variations in treatment integrity and outcomes is limited (Noell, Gresham, & Gansle, 2002). Some empirical evidence suggests program effectiveness can be associated with the fidelity with which it is implemented (Ruiz-Primo, 2006). For example, Gresham et al. (1993) found moderate, but significant correlations between percent of fidelity and treatment outcome measured by effect size, and found interventions implemented with higher fidelity were associated with larger effect sizes. However, Dane and Schneider (1998) concluded there was no consistent relationship between treatment outcomes and fidelity of implementation.

In a study aimed to directly measure the impact of intervention specificity, Mathes et al. (2005) compared the effectiveness of two intensive reading interventions that differed in theoretical orientation. One intervention, Proactive Reading, had a detailed scope and sequence and explicit daily lesson plans. The other, Responsive

Reading, relied on the teacher planning of instruction in response to student needs. These two interventions exemplify the two competing paradigms of curriculum implementation as explained by Cho (1998). Both interventions produced similar results on most outcome measures, although students receiving the more standardized intervention scored higher on one posttest. However, the instructors in this study were considered expert reading teachers and were matched to the intervention which best matched their philosophical view. Further, both interventions were implemented under highly controlled conditions, all which may limit the generalizability of these findings.

Similarly, in a study of first grade struggling readers McIntyre, Rightmyer, and Petrosko (2008) compared the effectiveness of a scripted reading model compared to four non-scripted reading models and found no significant differences for any measure of reading achievement. Wanzek and Vaughn's (2007) synthesis of studies corroborates the findings of both Mathes et al. (2005) and McIntyre et al. (2008). Wanzek and Vaughn concluded that interventions with differing levels of lesson standardization (i.e., standardized protocols or scripted lessons versus less standardized protocols, which allowed for more teacher decision-making in response to student need) resulted in no differences for student outcomes. These results suggest high levels standardization may not be required to achieve comparable results. A recent study conducted by Savage,

Carless and Erten (2009) found that students of teachers who made adaptations to a computer-based literacy intervention actually scored higher on several reading measures than the students of teachers who simply adopted the program as designed. However, only three teachers participated in this study, so it is unclear if the adaptations or some other teacher-related variable produced increased student achievement.

Other researchers have found a good deal of "indirect" success associated with problem-solving models, such as reducing the number of referrals in a school district or decreasing retention rates; however "direct" effects, such as improving the reading achievement of individual students, have been limited (Carney & Stiefel, 2008). Using meta-analysis, Burns, Appleton, and Stehouwer (2005) found field-based interventions using problem-solving approaches resulted in slightly lower student outcomes than interventions using standardized protocols; although systemic effects, such as the number of students referred to or placed in special education, favored the problem-solving approach. Similarly, Carney and Stiefel (2008) noted at one school, after 4 years of implementing a problem-solving model, 52% of students referred for academic problems were still receiving interventions, and 41% students were still considered at high or moderate risk.

Although the use of school personnel in many studies suggests the feasibility of implementation by real teachers in real schools, it is not clear if similar student outcomes would be achieved if teachers with less training and support implemented these interventions. If fidelity checks were not conducted, the level of teacher adherence to the intervention might differ. Because data involving student responsiveness to small group interventions in the RTI process may be considered in determining special education, it is important to consider the degree of fidelity needed to achieve valid, reliable outcomes. Clearly, it would be both impossible and imprudent to classify a student as a non-responder without knowing how faithfully the intervention was implemented.

Instructor qualifications. Several studies have examined the relationship between the qualifications of the person implementing the intervention. In a meta-analysis of one-to-one tutoring of children by adults, Elbaum et al. (2000) found that effect sizes for student outcomes had a significant relationship with the qualifications of the instructor. Across 31 studies, students tutored by college students made the greatest gains ($d = 1.65$), followed by students tutored by paraprofessionals ($d = 0.68$), teachers ($d = 0.36$), and community volunteers ($d = 0.26$) (Elbaum, et al. 2000). A recent meta-analysis of 97 studies on instruction for struggling readers, Slavin et al. (2011) found the largest effect sizes for students who received one-to-one tutoring by teachers ($d = .39$),

followed closely by paraprofessionals ($d = .38$), and then professional tutors or volunteers within programs with extensive training and structure ($d = .24$), and finally, volunteers ($d = .16$). Similarly, Brown, Morris, and Fields (2005) and Ehri, Dreyer, Flugman, and Gross (2007) both found stronger outcomes for children tutored by teachers than paraprofessionals using the same program, ($d = 0.47$ and $d = 0.52$, respectively), suggesting that the qualifications of the instructor may in fact play a role in the success of the intervention.

Spontaneous decision-making within interventions. Even within tightly scripted interventions, teachers must spontaneously assess the correctness of the response, form a hypothesis based on the student's response, and determine how to respond accordingly (Englert & Semmel, 1983). Effectiveness depends on the instructor's ability not only to impart knowledge or provide opportunities for learning and practice, but also to evaluate students' responses in order to strategically plan their next teaching move (Hattie & Timperly, 2007). These impromptu assessments and decisions made during instruction have been called interactive decisions because they take place while instructors are interacting with students (Borko & Shavelson, 1990). In their seminal review of instructional decision-making, Clark and Peterson (1986) estimated that teachers make between 0.5 to 0.7 decisions per minute, or approximately one decision

every 2 minutes. The highest percentage of these interactive decisions was in response to the learner, particularly when the participants were identified as struggling readers (Clark & Peterson, 1986; Semmel, 1977). Further, the most prominent types of interactions involved (a) teacher initiation, typically a question and (b) a response from the student, followed by (c) teacher feedback (Tharp & Gallimore, 1988). Research has clearly demonstrated the relationship between feedback and learning, as well as the differential effectiveness of various types of feedback and the manner in which it is given (Hattie & Timperly, 2007). The most powerful feedback is (a) focused on a clearly articulated goal; (b) simple, rather than complex, (c) unambiguously directed at the person it is intended for; (d) strategy-focused, rather than task-focused, and (e) not diluted with self-praise (i.e., "What a smart girl you are!").

Teacher behaviors and reading achievement. The National Reading Panel (2000), along with several researchers, identified feedback as a critical element of reading instruction (Duffy & McIntyre, 1982). The most effective teachers of reading comprehension are "reactive-corrective", monitoring student progress, recognizing when students have a misunderstanding, and responding with targeted instruction (Duffy & McIntyre, p. 17). Similarly, Garrison (1997) asserts "good teaching requires doing the right thing in the right way at the right time in response to problems posed by *particular*

people in *particular* places on *particular* occasions" (p. 271; emphasis added). Although teachers need to be able to make multiple rapid instructional decisions in order to react to students' reading comprehension errors, very little is known about how teachers respond to errors in reading comprehension (Meyer, 1986) and how those responses might impact student learning and achievement. With so few studies, the National Reading Panel (2000) called for future research on the teacher characteristics influencing reading comprehension, as well as greater attention to collecting data on the "comparability of instructors" and the need to "observe, document, and analyze" all aspects of instruction within reading comprehension interventions (p. 50).

Review of Empirical Studies

To investigate what is known about relationship between instructional behaviors during reading instruction and student reading achievement, I reviewed the existing empirical literature on the subject. Since literature focusing specifically on instructional behaviors in tutoring settings is limited, I expanded my review to include research conducted in general education or whole class settings. In the subsequent section, I will (a) describe the methods and criterion used to identify relevant literature for inclusion in this review, (b) present findings of these studies, and (c) offer a methodological critique of this body of literature.

Search Methods. To locate studies related to the impact of instructional behaviors on student achievement in reading, I conducted electronic, ancestral and forward searches. First, I conducted an electronic search for articles published in peer-reviewed journals using the Education Resources Information Center (ERIC), PsycInfo, Academic Search Premier and Education Research Complete databases. First, I entered combinations of the subject terms: "reading achievement" and "teacher characteristics", "teacher effectiveness", or "teacher-student interactions." This initial search resulted in 92 articles. I limited the search to include studies of school age children (6-12 years), and 36 articles remained. I reviewed the abstracts and only included studies (a) written in English, (b) which examined teacher's instructional behaviors related to reading instruction, and (c) which included a measure or description of student reading achievement. Seven articles met these three criteria.

In order to find additional articles, I conducted an ancestral search. By reviewing the reference section of each article, I identified relevant articles, and applied the same criterion for inclusion. I searched for additional articles written by authors of the identified studies. The ancestral search yielded an additional two studies. Finally, I used the Social Science Citation Index database to locate any related articles that cited the

articles identified for review. Using forward and ancestral searches, I located two additional articles that met the criterion for inclusion.

Results. Eleven studies were included in the final review. Participants were teachers and their students in grades spanning kindergarten through grade 8. Six studies examined teacher practices at one particular grade level, most frequently grade 2, while five studies involved students and teachers at multiple grade levels. The number of participating teachers ranged from 7 to 92, although one study (Samph, 1974) did not specify the number of participating teachers. All studies included in this review were conducted between 1974 and 2005. Ten of the eleven were published prior to 2004, when federal regulations changed to include RTI as a method of LD identification. Six studies used quantitative methods, two were purely qualitative in nature, and three studies utilized a mixed-methods approach. Researchers in nine of the studies conducted live observations of teachers' behaviors during literacy instruction. In one study, (Duffy, Roehler, & Radcliffe, 1986) teachers self-selected five lessons to audiotape and provided them to the researchers for analysis.

In the next section, I present the findings of the 11 studies and summarize what is known about the relationship between teacher behaviors and reading achievement.

Following that, I offer a methodological critique of this body of literature. Appendix A

includes a table summarizing the findings of the studies included in the literature review, including the design, purpose, participants, measures used to collect data on teachers and students, method of analysis, and results.

Content Review

Several researchers have studied the relationship between various instructional behaviors and student achievement. The teacher behaviors studied can be categorized in three ways: (a) preferred interaction style of teachers, (b) the quality and types of instructional talk used when the teacher is presenting new content, and (c) the manner in which teachers respond to students during instruction (i.e., instructional feedback in response to student errors).

Interaction style. Four studies considered how a teacher's interaction style impacted student outcomes in reading (Samph, 1974; Taylor, Pearson, Clark, & Wapole, 2000; Taylor, Pearson, Peterson, & Rodriguez, 2003; 2005). In the earliest study, Samph studied the impact of teaching style on student reading outcomes for students identified as low achievers. Following an unspecified number of observations, he defined two groups of teachers: (a) direct teachers, who were more authoritative, and used both critical statements and lecture more often, and (b) indirect teachers who used more praise and encouragement, and accepted the ideas and feelings of students more often. On a

norm-referenced measure of overall reading achievement, low achieving students assigned to indirect teachers outperformed their low achieving peers who were taught by a direct teacher. Additionally, those low achieving students taught by indirect teachers had more positive attitudes about school, leading Samph to conclude low achievers benefit both academically and emotionally when assigned to more accepting and encouraging teachers.

Later studies of teacher interaction styles focused on specific types of interactions and relationship with student reading outcomes. Taylor et al. (2000) observed 92 teachers in kindergarten through third grade classrooms over 5 months to examine the impact of the teacher's preferred interaction style on the reading achievement of their students. Using a checklist of effective instructional behaviors, two raters classified each of the teachers as most accomplished, moderately accomplished, or least accomplished. Observers, blind to condition, spent five 1-hour sessions watching instruction in these classrooms. Researchers coded for six interaction styles: (a) coaching/scaffolding, (b) modeling/demonstrating, (c) engaging students in recitation, (d) telling, (e) explaining procedures, and (f) discussion. Teachers identified as most accomplished engaged in less whole-class instruction, and instead provided instruction in small groups. Students in classrooms with accomplished teachers were on task substantially more of the time. In

terms of interaction style, the most accomplished teachers used coaching (48%) more than their colleagues who were rated moderately accomplished (21%) or least accomplished (2%). Further, the least accomplished teachers told their students the correct answer significantly more often (75%) than those teachers rated as moderately accomplished (38%) or most accomplished (7%). By the end of the school year, first grade students assigned to teachers rated moderately accomplished or accomplished read on average 19 more words correct per minute (wcpm) than the students with teachers rated least accomplished. Similar differences in fluency were apparent in second grade as well; classes with accomplished teachers read on average 10 more wcpm than teachers rated least accomplished, despite an initial 8 wcpm advantage for students assigned to the least accomplished teachers. Differences in fluency were not present in third grade.

In a follow up study, Taylor and her colleagues (2003) used HLM to investigate the influence of teacher behaviors on the reading scores of students in grades 1 through 5 living in high-poverty areas. Research staff observed each teacher three times throughout the school year, coding for 19 behaviors. They found that 35% of the variation in spring fluency scores in first grade could be attributed to teacher differences specifically differences in the number of higher-level questions teachers asked. In grades 2 through 5, 46% of variance in fluency scores occurred between teachers. For each standard

deviation above the mean in coaching, fluency increased by 4.2 wcpm. Every standard deviation (SD) increase in reading practice related to an average gain of 3 wcpm. Similar results were found for comprehension scores. Also, for each SD increase in the time students spent on task, comprehension scores increased by 2 points. More time spent teaching comprehension skills, as opposed to strategies, and more passive responding was negatively related to comprehension achievement. For each SD increase in time spent in comprehension skill instruction, students' comprehension scores decreased 2.4 points. Teacher level variables accounted for 35% of variance in spring reading comprehension scores, 27% of which was attributed to asking more higher-level questions. Forty eight percent of the variance in comprehension scores was attributed to teacher differences. Again, asking a greater number of higher-level questions was related to increased student reading achievement.

In a later study Taylor et al. (2005) found teacher variance accounted for significantly more variance than school level variables in comprehension and writing scores. Variance in comprehension outcomes could be attributed to teacher-level variables (24%), more than double than that found between schools (10%). In fluency, however, teacher variance was more comparable to school variance (19% and 22%, respectively). Similar to the previous study, the number of higher-level questions asked

by teachers was related to increases in student reading fluency. Teacher impact on writing (32% of variance explained between teachers) was far greater than the amount of variance attributed to schools (4%), indicating that teachers may exert a greater influence over writing and reading comprehension than which school a student attends. Students whose teachers engaged in coaching made greater growth in writing. Similar to the findings of the 2003 study, instruction in rote comprehension skills was negatively related to both comprehension and fluency growth.

Findings from these four studies demonstrate a relationship between the type and style of teacher interactions and differential student reading outcomes. Higher student outcomes were related to teachers who frequently used coaching, engaged in more small group instruction, and who asked a greater number of higher-level questions. Increases in on task behavior and higher levels of praise and encouragement were also related to improved student outcomes. In contrast, lower levels of student achievement were associated with teachers who engage in more whole group instruction and lecture, told students the correct answer rather than coaching them toward finding it themselves, and focused their teaching on rote skills or behaviors, rather than strategies (i.e., metacognitive plans and processes). Some of the variance in student reading scores can

be attributed to the types and quality of instructional interactions, suggesting teacher differences may impact student responsiveness within RTI.

Only one (Samph, 1974) of the four studies specifically examined the impact of teacher interaction style for students identified as low achieving. Yet, Samph did not look at specific teacher behaviors, but instead separated teachers into two groups based on a high-inference coding system. All studies were conducted in whole-group general education settings. It is unclear if the impact of interaction style is different for students who are struggling in reading and receiving small-group intervention.

Instructional quality. Five studies focused on the relationship between the quality of instruction and student learning. In a study of seven grade 5 teachers, Duffy, Roehler, and Rackliffe (1987) observed teachers during comprehension instruction. A researcher-developed instrument was used to measure the explicitness of the teachers' explanations. After the lesson concluded, Duffy and his colleagues interviewed the five lowest students (as nominated by the teachers) and asked them to tell everything they remembered about the lesson. The researchers prompted the students by asking (a) what was taught? (b) when would it be used? and (c) how would you use it? The level of teacher explicitness had a moderate, significant correlation ($r = 0.59$) with student understanding of what was taught. Although all teachers were teaching the same content

– how to use context clues when reading – teachers whose students scored lowest presented the information as set steps or rules, as opposed to modeling the metacognitive processes involved in using context clues while reading. Teachers who provided explicit explanations were noted to engage in modeling and think alouds during instruction, which related to improved learning outcomes for struggling students.

Similarly, Mariage (1995) examined the instructional dialogue of teachers during comprehension instruction and the mean number of ideas recalled by students with learning disabilities. After testing students across 15 classrooms, he identified three teachers whose classes made the most gains, and three teachers whose students experienced the least amount of growth. Similar to the findings of Duffy et. al (1986) the high gaining (HG) teachers used more modeling statements and demonstrated "more thinking in action" (Mariage, 1995, p.223). The teachers whose students made the least progress, identified as low gaining (LG), used more statements to maintain classroom management and double the number of evaluative statements than the HG teachers (48 vs. 24). More often than the LG teachers, the HG teachers turned the conversation over to the students and provided more opportunities for students to extend their thinking.

Findings from a study conducted by Lara and Medley (1987) further support students who have learning disabilities or are low-achieving benefit from explicit

instruction and teacher-directed modeling. To determine which teacher behaviors are most effective in improving student achievement gains for students of both high and low ability levels, Lara and Medley observed teachers in grades 3 through 8 and counted the frequency of behaviors believed to be indicators of teacher competence. These behaviors included: (a) using a variety of methods in presenting material, (b) establishing one-on-one relationships with students, and (c) providing daily opportunities for success. No single pattern of behavior emerged as best for all learners. High achieving students made the most gains when taught by a teacher who demonstrated proper listening skills and non-verbal communication skills, and avoided the use of praise or rewards. Lower achieving students were most successful when paired with teachers who gave clear, explicit directions and did not allow other individuals to speak. This last finding may indicate that for lower achieving students, direct instruction from the teacher is most beneficial. It may also be that lower achieving students do best when their attention is directed at just one speaker, the teacher.

Wharton-McDonald, Pressley, and Hampston (1998) studied six first grade teachers whose students made the most significant gains in literacy and teachers who were not as effective. Through observation and interviews, they defined eight characteristics that distinguished the high-achievement teachers from their colleagues: (a)

balancing explicit decoding instruction with authentic reading and writing activities; (b) pursuing multiple instructional goals throughout the day (taking advantage of teachable moments/mini lessons); (c) use of scaffolding; (d) encouraging self-regulation through metacognitive modeling; (d) integrating reading and writing activities; (e) demonstrating high expectations for all students; (f) having exceptional classroom management and well-established classroom routines; (g) being well-prepared and extremely organized, and (h) exhibiting a strong sense of purpose and intentionality.

While the majority of researchers have sought to identify the behaviors of the most effective teachers through observation, an early experimental study was conducted by Alpert (1975) to determine if manipulating teacher behaviors would improve student achievement. She hypothesized that an increase in "good" instructional practices would improve reading scores for low achieving readers. Consulting with teachers and reading specialists, Alpert developed a list of 26 "good" teacher behaviors, including (a) increasing the time struggling students received small group reading instruction (b) teaching the lowest group at the time of day the teacher felt most motivated, (c) using a greater number of materials to teach reading, (d) working with fewer students at a time, and (e) increasing selected verbal behaviors. Preferred verbal behaviors included praise, reinforcement, encouragement, and demonstrations. Random assignment occurred at the

school level, resulting in nine teachers in the control condition, and eight teachers who were instructed to employ the "good behaviors" when working with their lowest group of readers for a period of 11 weeks. Results indicated no post-test differences on vocabulary or comprehension measures between the two conditions. The author concluded no evidence supported the impact of teacher behavior and student performance, however methodological limitations impede the strength of this conclusion.

Three of the five studies focused on low achieving students and had similar findings (Duffy et al. 1986; Lara & Medley, 1987; Mariage, 1995). Results suggest low performing students benefit from (a) explicit explanations (Duffy et al., 1986; Lara & Medley, 1987), and (b) modeling accompanied by verbalizing the accompanying cognitive processes (Duffy et al., 1986; Mariage, 1995). Similarly, Wharton-McDonald and colleagues (1998) found the teachers whose classes that made substantial literacy growth also demonstrated similar behaviors, as well as having masterful classroom management skills. However, like the studies on interaction style, these studies were conducted within general education whole-class contexts. Although some teacher behaviors have been linked to increases in student achievement, the individual rates of student growth when assigned to teachers with different behavior patterns have yet to be studied.

Responding to students. Two of the reviewed studies focused on teachers' responses to students and the type of feedback provided to individual students during class discussions. To determine the relationship between type of student mistake, teacher response, and student achievement, Hoffman et al. (1984) asked 22 teachers to self-select and audiotape five guided reading lessons. The type of reading errors students made while reading were coded, as were the teachers' responses to miscues. Teacher feedback was designated as either (a) terminal, where the teacher told the student the word or called on another student; (b) sustaining, when the teacher encouraged the student to correct their own miscue; (c) attending, when the teacher let the student know an error was made, but did not offer or request a correction; (d) grapho-phonetic, when teacher prompted the student to attend to phonic elements; or (e) contextual, if the teacher prompted the student to pay attention to meaning or surrounding words. Results suggest that there is indeed a relationship between the type of error made by the student and the teacher's choice of response. Teachers were most likely to interrupt a student during oral reading if the miscue changed the meaning of what was being read. Results of the regression analysis indicated that increases in terminal feedback were related to decreases in student-reading growth and increases in the number of hesitations students made during oral reading. Higher levels of terminal feedback and increased error rate had a

significant negative impact on student achievement, which suggests students should be reading texts in which they make few errors and when errors occur, provide opportunities for self-correction, rather than supplying the correct word.

Similarly, Martin, Veldman, and Anderson (1980) found a negative relationship between terminal feedback and student achievement. Researchers observed six first grade teachers and noted how teachers responded to both correct and incorrect answers. Students who received more terminal feedback and high rates of criticism scored lower on the posttest. Students were most successful when they had received more sustaining feedback, were asked more comprehension questions, and were provided opportunities to self-correct errors. Interestingly, students who were most frequently called on by the teacher performed less well than students who were called on less often. Behaviors did not vary significantly by teacher, but had significant variation at the student level, indicating individual teacher-student interactions may be more influential than a teacher's overall style.

Findings from both studies suggest teachers should avoid the use of terminal feedback, and instead provide sustained feedback more frequently. Similar to the other studies in this review, data were collected within general education classrooms, making it difficult to generalize the results to other contexts. Hoffman and colleagues (1984)

examined if teacher feedback differed for high and low achievers, but found that feedback was related to the type of error made by a student, rather than perceptions of aptitude.

Methodological Review

Researchers have studied the relationship between teachers' instructional behaviors and student achievement using three general methodological approaches. In this review (a) six studies used a quantitative approach, (b) two sought to describe the relationship using qualitative methods, and (c) three utilized a mixed-methods approach. Each methodological approach has different indicators of quality. What follows is a review of the methodological rigor of the 11 studies, organized by general methodology.

Quantitative studies. Of the 11 studies reviewed, 6 used quantitative methods. Four studies were non-experimental in design. In these studies, researchers collected quantitative data to investigate the relationship (i.e. correlation) between instructional behaviors and student achievement, but participants were not randomly assigned to treatment conditions the studies (Hoffman, et al., 1984; Lara & Medley, 1987; Samph, 1974; Taylor et al., 2000). The researchers in the other two quantitative studies (Alpert, 1975; Martin et al., 1980) used an experimental approach, assigning teachers to

treatment or control conditions in order to determine if increases in specific behaviors would improve student literacy outcomes.

The researchers in the four non-experimental studies used data analytic methods of multiple regression, and analysis of variance (ANOVA) or covariance (ANCOVA) techniques. When considering the results of studies using correlations, one must consider the selection and description of participants (Gersten et al., 2005), the reliability and validity of the measures used, and the fit of the analytical approach (Thompson, Diamond, McWilliam, Snyder & Snyder, 2005). These three considerations influence the strength of the findings, as well as the generalizability of the results.

Two studies in this review had an experimental design (Alpert, 1975; Martin et al., 1980). Quality experimental studies should include (a) the use of sampling procedures to ensure comparability across conditions, (b) a detailed description of the intervention, (c) procedures for measuring fidelity of implementation, and (d) appropriate data analytic techniques (Gersten et al., 2005). When considering the results of experimental studies, any threats to validity must be taken into account. Such threats include differential selection of participants, instrumentation, controlling for extraneous variables, and participant reactivity (Gay, Mills, & Airasian, 2006).

Sampling. Results for both non-experimental and experimental studies should be interpreted within the particular context in which the study was conducted. Methods for selecting participants and characteristics of both the setting and participants impact findings, and thus should be thoroughly described by authors. Readers must consider these factors related to sampling to determine the validity and generalizability of the results presented by the authors.

Teacher selection. Results should be interpreted within the context of how teachers were selected for participation. Most of studies reviewed provided limited information on teacher selection processes, the percentage of teachers who agreed to participate, or attrition rates. All studies failed to address the issue of selection bias, that is, if the teaching practices of participating teachers differed from those who did not wish to participate. Although it is unclear if teachers who volunteer to be observed tend to be more accomplished teachers, those teachers who volunteer to be observed by researchers may be more confident in their teaching abilities. For example, Martin et al. (1980) indicated all teachers in their study volunteered to participate. When considering the results of studies whose participants included only volunteers, the results can only generalize to other teachers who would volunteer to be observed. It is unclear if participating teachers were representative of all teachers in the district or if non-

participators differed from participators. Further, no study provided information regarding attrition.

The number of teachers observed in these quantitative investigations ranged from 15 (Lara & Medley, 1987 and Martin et al., 1980) to 92 (Taylor et al., 2000). Samph (1974) did not provide the number of teachers who participated in his study. Lara and Medley selected teachers based on student growth on standardized test results, identifying teachers whose students made high gains and those teachers whose students made less substantial gains to determine the teaching practices that distinguished effective teachers from those who were less effective. Taylor and her colleagues (2000) took a different approach, intending to only examine the practices of exemplary teachers. They asked principals to recommend exemplary teachers for inclusion in their study. However, in order to verify the teachers were, in fact, exemplary, the researchers examined student achievement data and scores on researcher-designed measures; and determined not all of the nominated teachers were exemplary. The final analysis was based on a wide variation in teacher practices and provided information about both accomplished and less accomplished teachers.

Setting and Student Characteristics. All researchers examined teacher practices within general education classrooms. Two studies (Alpert, 1975; Samph, 1974) examined

the impact of teacher behaviors on low achieving students, while two others (Hoffman et al., 1984; Lara & Medley, 1987), determined if there was differential impact of teacher behaviors on high- and low-achieving students. However, no study provided evidence beyond teacher or school district identification confirming the achievement status of the students, as recommended by Gersten et al. (2005). For example, both Alpert and Hoffman et al. selected students based on their classroom assignment to reading groups; while Samph selected participants based on standardized test scores two or more grade levels below the national norm, as well as presenting the mean IQ score for the sample (109.2). Taylor et al. (2000) used stratified random sampling to identify equal numbers of low, average- and high-performing students in each class, based on teachers' recommendations of performance, however they did not identify the association between gains for these specific student groups and teacher behaviors. Using labels provided by teachers or schools calls to question the validity of sample selection, as true levels of aptitude or achievement were not independently defined and confirmed (Gersten et al., 2005). Without adequate information about the student population, it is difficult to determine the relevant population for which findings could be generalized. In fact, demographic information on the students was limited in all studies. No study provided information on the race, ethnicity, or socioeconomic status of the student participants.

These contextual variables have been shown to impact achievement (Bradley & Corwyn, 2002). In all studies, measures of teacher behavior were correlated to only a sampling of student achievement scores within each class due to limited resources. This body of research would be more methodologically sound if demographic data were provided for both the student population, as well as the participating teachers, in addition to including data for all students receiving instruction in the analysis.

Measurement. Researchers collected data on both teachers and students. Data on teacher behaviors was collected primarily through observation. Threats to validity and reliability of observational data involve the observation procedures, specificity of coding, the number and length of observations conducted, and inter-rater reliability. Data on student reading achievement was collected using norm-referenced, informal, and curriculum-based measures. In the subsequent sections, I present the methodological strengths and limitations in the measurement of both teacher and student variables.

Teacher Measures. Live observation data was used in four of the six quantitative studies. Alpert (1975) and Hoffman et al. (1984) did not conduct live observation, instead Alpert provided teachers in the treatment condition with a list of behaviors to engage in throughout intervention, relying only on self-report to assess fidelity of implementation. Hoffman et al. collected data by asking teachers to self-select five lessons to audio

record. The four studies that used data from live observations included researcher-developed observation instruments comprised of detailed observational field notes and coding for specific teacher behaviors. The focus of the instrument was guided by the research question(s) of interest. The level of detail provided concerning the development and validation of the instrument utilized varied across studies. Taylor and her colleagues (2000) provided the most extensive information on the development and validity of the observation instrument. Only one study (Taylor et al., 2000) included descriptions of the observation procedures in adequate detail to allow for replication, as the observation protocol was included in the appendices. Researchers measured a variety of teacher behaviors through observation, including response to student errors (Hoffman et al., 1984), behaviors thought to be indicative of teaching competence (Lara & Medley, 1987), and interaction type (Taylor et al., 2000).

When considering the results of a study, it is prudent to examine both the reliability and validity of the measures used, as any threats may attenuate the findings. Two studies included sufficient detail, examples, and the operational definitions of the teacher behaviors coded for during classroom observations (Hoffman et al., 1984; Taylor et al., 2000). Other researchers were more ambiguous in their descriptions, increasing the possibility of observer bias. For example, it is unclear how observers determined how to

code "accepts feelings" (Samph, 1974) or "maintains self-control" (Lara & Medley, 1987), since descriptions or illustrations of the codes were not included. Alpert (1975) found no significant relationship between teachers in the control and treatment conditions, however this is attenuated by the lack of data to formally assess adherence to the intervention. Without fidelity of implementation data, it cannot be ascertained how frequently, if at all, teachers demonstrated the desired behaviors.

One indicator of quality within correlational studies is the reporting and analysis of reliability coefficients for all measured variables (Thompson et al., 2005). Three studies reported on interater or intercoder reliability, and all reported initial agreement greater than 80%. None of the investigators provided the formula used to calculate percent agreement for their study. Three of the studies included descriptions of observer training and established a criterion of agreement to be reached (typically 80%) before observers were allowed to conduct classroom observations (Hoffman et al, 1984; Martin et al., 1980; Taylor et al., 2000). In only one study (Taylor et al., 2000) random checks of interrater reliability were conducted to guard against observer drift.

The number of observations per teacher varied across studies. Taylor et al. (2000) observed each teacher conduct entire literacy lessons five times over the course of a school year, while Lara and Medley (1987) observed teachers between 5 and 10 minutes

on six occasions. Samph (1974) did not specify the number of observations he conducted. Results of studies that relied on a single observation must be considered carefully, as the observation was simply a snapshot of classroom practice and may not accurately represent the daily instructional behaviors of the teacher. To get a more valid picture of daily classroom events for each teacher, multiple observations should be conducted across different times of the school day, and year.

Additionally, when designing studies using observational data, researchers should consider the impact of observer effect, where persons being observed behaves atypically because they are being observed. Gay et al. (2006) suggest that observer effects typically decrease over time. Martin and his colleagues (1980) observed each teacher 15-20 times over one school year, the greatest number of observations per teachers for any study reviewed, drastically reducing the threat of observer effect. In contrast, both Hoffman et al. (1984) and Taylor et al. (2000) collected observational data for five lessons per teacher, while Lara and Medley (1987) observed their teachers on six occasions for only several minutes each time. Without this information, the reliability of the observational data collected is called into question.

Student Measures. Reading comprehension was the most frequently measured construct of student reading in this body of literature. One study used a standardized

measure of reading comprehension, the Gates MacGinitie (Alpert, 1975), while another (Taylor et al., 2000) used an informal measure of comprehension, by tallying the number of ideas a student recalled after reading. Four studies used norm-referenced standardized tests of overall reading achievement, such as the Iowa Test of Basic Skills (Lara & Medley, 1987), the California Achievement Test (Hoffman et al., 1984), and the Metropolitan Achievement Test (Martin et al., 1980; Samph, 1974). Taylor et al. (2000) measured reading achievement using a combination of curriculum-based (CBM) and norm-referenced measures of reading. CBM has been shown to be more sensitive to reading growth than traditional, norm-referenced measures (Deno, Fuchs, Marston, & Shin, 2001). Although Thompson et al. (2005) suggested that all correlational studies should report reliability coefficients for all variables, reliability coefficients were not reported for any of the student measures in any of the studies.

Statistical Treatment. One study used hierarchical linear modeling (HLM) because of the nested nature of the data. This is appropriate to partition the variance that can be attributed to student-level, teacher-level, and school-level variables. Taylor et al. (2000) built a two-level model to predict student outcomes based on student and teacher characteristics, due to insufficient power to detect school-level effects. The analysis appropriately controlled for confounding variables, such as classroom composition

variables, (e.g., socioeconomic status, percent minority, and initial mean classroom achievement) to attempt to account for differences across classrooms that should not be attributed to teacher differences. Taylor et al. (2000) also conducted a multivariate analysis of variance (MANOVA) to examine the relation between school effectiveness based on student literacy achievement and classroom instruction. After the MANOVA indicated group differences existed, several post hoc ANOVAs were conducted to investigate the differences. To minimize the likelihood of Type I error, the authors used Tukey post hoc tests, although they did not specify which version of the Tukey post hoc test they employed, as suggested by Huck (2008).

Summary. The six quantitative studies presented in this review reveal some consistent methodological concerns. Only one study (Taylor et al., 2000), met most, but not all, of the quality indicators put forth by Gersten et al. (2005) and Thompson et al. (2005). No study provided comprehensive information on the selection of teachers and students, and relevant demographic information. This lack of information impedes interpretation of findings, as the context in which the study took place is unclear. Although two studies focused on low-achieving students in particular, the method used to identify the students as low achieving was of questionable validity. All studies were conducted within general education settings, and so the quality and impact of

instructional interactions within small group reading interventions remain unknown, particularly for students validly identified as at risk for reading failure. Further, over half of studies used norm-referenced measures, which are less sensitive to short-term growth (Deno et al., 2001). Future studies should present information regarding reliability coefficients for the measures used, the specific methods for calculating inter-rater reliability of observation coding, and include the achievement data for all students receiving instruction, not just a subgroup. Further, the impact of teacher behaviors on student reading achievement must be considered within small group intervention contexts, as the student-teacher ratios are smaller, possibly concentrating the influence of teacher behaviors.

Qualitative studies. Two of the reviewed studies employed qualitative methods (Duffy et al., 1986; Wharton-McDonald et al., 1998). In considering the results of qualitative observational studies, components of the design and the approach to collecting and analyzing data impact the perceived credibility and trustworthiness of the study (Brantlinger, Jimenez, Klingner, Pugach, & Richardson, 2005; Creswell, 2007). What follows is an analysis of the qualitative studies in this review using evaluation criteria proposed by Brantlinger and colleagues.

Design considerations. When evaluating qualitative studies using observational data, the following aspects of design should be considered (a) the number and duration of observations, (b) the impact of the research on the setting, (c) the appropriateness of the setting and/or participants, and (d) the data collection procedures (Brantlinger et al., 2005). The first two, the number and duration of observations and the impact of research on the setting, address one threat to the validity of observational studies, known as observer effect (Gay et al., 2006). Observer effect is not a threat unique to qualitative studies, as any study based on observational data is susceptible to this threat. If the participants alter their behaviors in the presence of the observer, the researchers may draw inaccurate conclusions. Prolonged engagement in the setting can guard against this threat, as participants become more comfortable with the presence of observers over time and data saturation occurs. The researchers in the Wharton-McDonald et al. (1998) study observed each teacher for approximately 12 hours, the longest of any of the 10 studies reviewed which collected data through observation, thus minimizing the impact of observer effect. Although Duffy et al. (1986) based their analysis on one lesson observation from each classroom, describing typical practices in these classrooms was not the focus of their study. Instead, researchers sought to describe the type and

explicitness of instructional language used during one lesson and its impact on student understanding, so observer effect was not of particular concern for this study.

Another threat to the validity of observational studies is observer bias, which can be remedied through triangulation by confirming findings across multiple data sources (Gay et al., 2006). Both qualitative studies reviewed included various sources of data. For example, Wharton-McDonald and her colleagues (1998) conducted observations and interviews, in addition to analyzing multiple classroom artifacts, such as books, student writing, and other materials. To describe the quality of instructional talk, Duffy et al. (1986) transcribed several lessons from each teacher and used a rating scale to score teacher explanations. This information, combined with data collected from structured interviews with students to gauge their understanding of lesson content, as well as ratings of the quality of student understanding, allowed the researchers to triangulate the findings across all data sources.

In considering the appropriateness of the participants and setting, Wharton-McDonald and colleagues (1998) took measures to ensure that the teachers they were observing were, in fact, outstanding or average as identified by their literacy coordinators. The researchers independently measured student achievement to be sure they had accurately identified the highly effective teachers from their less effective

colleagues, in order to describe differences in their practices. Duffy et al. (1986) provided rich description of the participating teachers, allowing readers to gauge the degree of transferability to their own contexts.

Data collection. A thorough description of the data collection and analytic procedures support the reliability of the information presented (Creswell, 2007). For instance, in the study conducted by Duffy et al. (1986) the lessons were audiotaped and transcribed verbatim, then scored by trained raters for explicitness of teacher response. Likewise, the student interviews that followed the lesson were audio taped and transcribed, and scored by two independent raters to measure student understanding. Further, Duffy et al. reported intercoder reliability of .90 for explanation ratings and .83 for student awareness ratings. In contrast, the lessons in the study by Wharton McDonald and colleagues (1998) were not audiotaped and transcribed verbatim, instead the researchers relied on field notes taken by two observers simultaneously that as closely as possible recorded the language of the teachers and students verbatim. Further, the authors in this study did not engage in independent coding and comparison, and presented no information regarding intercoder reliability.

Data analysis. During data analysis, researchers can take several measures to support the credibility of their findings. For instance, data triangulation, the use of

external auditors, conducting member checks or peer debriefing, and establishing an audit trail, are recommended when relevant to the study at hand (Bratlinger et al., 2005).

Additionally, Bratlinger and her colleagues (2005) suggest that high quality qualitative research consists of (a) sorting and coding the data in a meaningful way, (b) a rationale for information included or omitted, (c) documentation of measures taken to establish credibility and trustworthiness, (d) disclosure and reflection of the researchers personal perspectives, and (e) sufficient use of quotations and presentation of evidence from observations, as well as (f) evidence of document inspection and (g) discussion of the relationship of the findings to related research. The reviewed qualitative studies included most of these components, however, neither study included information regarding researcher reflectivity, such as disclosing any biases, values, or experiences that may have influenced their interpretation of the data. Further, neither study employed external auditors to scrutinize the data and substantiate their conclusions. Wharton-McDonald et al. (1998) conducted member checks; Duffy et al. (1986) did not. Had these credibility measures been employed, the strength of the results would be more robust.

Summary. Overall, the methodological rigor of the qualitative studies in this review was mostly sound. Both studies included a combination of procedures to ensure credibility and trustworthiness. However, neither study used external auditors to verify

that the conclusions of the researchers fell in line with those provided through independent examination of the data from people not affiliated with the study. When comparing the reliability of the two studies reviewed, the data collection procedures used by Duffy et al. (1986) are favored. By audio recording and transcribing all lessons verbatim the researchers provided an additional measure of credibility, as more accurate data was coded, transcripts allow for multiple people to analyze the same raw data, and calculation of interrater or intercoder reliability is possible. Additionally, verbatim transcripts can be shared with external auditors, as well as presented to participants during member checks.

Mixed methods studies. Mixed methods integrate qualitative and quantitative approaches by combining the techniques, language, and concepts associated with each class of research (Anfara, 2006). Advantages to using a mixed approach include counterbalancing weaknesses associated with each approach, as well as the potential asset of presenting converging findings (Johnson & Onwuegbuzie, 2004). Research conducted using mixed methods may take various forms, complicating the establishment of a single set of criterion by which to evaluate the quality of mixed methods studies. When evaluating studies employing mixed methods, Creswell (2005) has suggested consideration of: (a) identification and rationale for the selected design, (b) specification

of the prioritized approach, (d) appropriateness of the approach in relation to the research questions, and (e) clear explanation of data analytic strategies, including sequencing of approach. These criteria will be used to evaluate the three mixed-methods studies in this review.

Design, rationale, and methodological priority. None of the studies stated upfront that they were employing mixed methods. For example, Taylor et al. (2003) took a primarily quantitative approach using HLM to determine the effect of teaching variables on student achievement. However, following the results of the HLM, the authors included descriptions, examples, and counterexamples to compliment their quantitative findings. The authors provided the following rationale for including qualitative data: "To get a picture of what the results look like in everyday practice [and] offer a clearer sense of what was going on in these classrooms " (p. 19). This is an example of one of the advantages of mixed methods, as the combination can provide insights that might be missed with the use of a single method (Anfara, 2006). Similarly, the study conducted by Taylor et al. (2005) appeared to be primarily quantitative in nature; however a closer examination of the methods revealed that Taylor et al. employed both quantitative and qualitative methods in collecting and coding observational data. Similar to the 2003 study, qualitative illustrations supported the quantitative findings.

However Taylor et al. (2005) did not provide a rationale for taking a mixed methods approach, although it can be inferred that the qualitative data was collected to explain the quantitative findings. Currently, this explanatory design (Creswell, 1994) is the most prevalent use of mixed-methods within educational research (Anfara).

While both studies conducted by Taylor and colleagues (2003; 2005) used an explanatory design, Mariage (1995) used an explanatory mixed-methods design, which Creswell (1994) explains is when qualitative data is collected to investigate a phenomenon, and then quantitative data is used to explain it. Mariage did not present an explicit rationale for employing mixed methods, although he argued that previous research had used static measures of teacher behavior that lacked the flexibility needed considering the sociocultural complexities within classroom discourse.

Research questions. The use of mixed methods in these studies was appropriate given the research questions each aimed to address. Mariage (1995) sought to understand how teachers use dialogue differentially to assist their students in construction of meaning during reading comprehension instruction. Additionally, he sought to determine if the instructional language used by teachers whose students made more substantial gains differed from those teachers whose students made smaller gains. To explore this phenomenon, he transcribed, coded, and analyzed audiotapes of lessons from teachers in

both groups. He calculated the frequency of six types of teacher statements and presented descriptive statistics for each group. Although no tests of statistical significance were conducted, the differences in the nature of teacher behaviors suggest practical, if not statistical, significance. Using descriptive statistics in this manner has been referred to as a logic-based approach to correlational research (Thompson et al., 2005). The method used was fitting to answer the research questions in this study.

Similarly, the studies conducted by Taylor and colleagues (2003; 2005) sought to answer a similar question, to determine how specific teaching variables impact student achievement. Since the currency of classroom instruction is instructional language, both studies conducted observations and quantified instructional discourse, which was later used in building statistical models to determine the impact of teacher behaviors on student achievement. The subsequent illustrations of these behaviors presented through description and quotations, demonstrate how qualitative data can complement quantitative findings using one to bolster the other.

Explanation and sequence of data analysis. All three studies provided detailed description of the sequence of procedures used to collect and analyze data. All researchers simultaneously gathered and analyzed qualitative and quantitative data. Explanations in all three studies were sufficiently detailed to allow for replication, as

coding schemes and rubrics were included for each of the studies. The coding scheme used in the two studies conducted by Taylor et al. (2003; 2005) was devised and refined in previous studies, and thus was more explicit. In comparison, Mariage (1995) developed his own coding system based only on the data collected in this small-scale study, so the coding scheme was less developed – the categories of codes were more broad and descriptions of codes less detailed.

Summary. The methodological limitations of the mixed-methods studies in this review provide directions for future mixed-methods research in this area. None of the three studies explicitly identified their design as mixed-methods. Of the three mixed-method studies reviewed, two prioritized the quantitative approach and one gave precedence to qualitative methods. Reviewing these studies emphasizes the importance of using a comprehensive and coding scheme, such as those used by Taylor and her colleagues (2003; 2005) even if it may evolve as data is collected, as advocated by Mariage (1995) who prefers the flexibility of altering and expanding codes in response to contextual variations.

Methodological Considerations for Future Research

The methodological strengths and limitations of the studies in this review provide directions for future research in this area. Only 2 of the 11 studies provided

comprehensive information on the selection of teachers and students, and relevant demographic information (Taylor et al., 2003; 2005). The method used to identify the students as low achieving in all 11 studies was of questionable validity. Future research should provide comprehensive information on the level of functioning of those students labeled "at-risk", as recommended by Gersten et al. (2005). Further, the reliability coefficients for the measures used should be presented and accompanied by adequate analysis concerning the appropriateness of those measures for use in the study at hand, as recommended by Thompson et al. (2005). Limitations of the studies reviewed should be addressed in future studies, such as the use of audio-recording and verbatim transcription of the lessons, clear explanations of the calculation of interrater or intercoder reliability, as well as the use of external auditors.

Synthesis and Directions for Future Research

Current knowledge. Although the 11 studies included in this review were derived from various methodological perspectives, overall findings suggest certain teacher behaviors are related with superior student outcomes in reading. Three critical behaviors appeared to be associated with higher student outcomes, supported by the findings of multiple studies: (a) explicitness (Duffy et al., 1986; Hoffman et al. 1984; Lara & Medley, 1987; and Wharton-McDonald et al. 1998); (b) metacognitive modeling

(Duffy et al., 1986; Mariage, 1995; and Wharton-McDonald et al., 1998); and (c) use of scaffolding (Mariage, 1995; Taylor et al., 2005; Wharton-McDonald et al., 1998)

Similarly, two behaviors consistently appeared related to inferior student outcomes across multiple studies: (a) use of terminal feedback (Hoffman et al., 1984; Mariage, 1995; Taylor et al., 2000) and (b) using lecture to present the use of rote skills (Duffy et al., 1986; Taylor et al., 2003; 2005). Positive student outcomes were associated with the following teacher behaviors in at least two studies: (a) coaching (Taylor et al., 2000; 2005) (b) asking more higher-level questions (Taylor et al., 2003; 2005), (c) excellent classroom management (Taylor et al., 2000; Wharton-McDonald et al., 1998); and (d) providing sustaining feedback more frequently (Mariage, 1995; Martin et al., 1980). The studies conducted by Taylor and her colleagues (2003; 2005) found that some variance in student reading scores could be attributed to the types and quality of instructional interactions, suggesting teacher differences within tutoring settings may, in fact, impact student responsiveness within RTI. All but one of these studies (Alpert, 1975), suggest teachers' instructional behaviors are related to student reading scores.

Directions for Future Research. Although several studies focused on low achieving students or compared the impact of teacher behaviors on low and high achievers, each of these studies have examined the relationship of instructional behaviors

within a whole class setting. Quite simply, no study was conducted within the context of special education or remedial interventions. Within smaller groups, the number of student-teacher interactions increase (Ozerk, 2001), and it is unclear if increased opportunities for interaction strengthen the influence of teacher behaviors. Future research should look at the impact of teacher behaviors specifically within small group settings. Of particular interest is how these instructional differences might impact student responsiveness or growth within the context of RTI. More research is necessary to determine if feedback, instructional language, and overall interaction style of teachers or tutors in small group interventions would similarly relate to student achievement.

Chapter III: Method

Purpose of the Study

This study examines the types of instructional behaviors tutors exhibit while interacting with students during the implementation of a scripted reading intervention.

The research questions of interest are:

1. What types of spontaneous instructional behaviors do tutors exhibit in reaction to student oral responses within a scripted reading comprehension intervention?
2. How do tutors differ from each other in regard to the type, variety, and frequency of the identified spontaneous instructional behaviors?
3. How do tutors differ from each other in terms of their fidelity of implementation of the scripted lesson plan?
4. What is the relationship between the spontaneous instructional behaviors of tutors and student reading comprehension outcomes, controlling for student demographic variables, prior achievement, and fidelity of implementation?

The first research question was addressed using a primarily qualitative approach. The second and third were answered using both qualitative and quantitative methods. The fourth question was answered using a quantitative analysis of the frequency of codes

created to answer the first research question and student reading comprehension outcomes following the intervention.

Context

This study is situated within a larger intervention study. Using a validated screening procedure Speece et al. (2010) identified two consecutive cohorts of students with a relatively high probability of reading failure using a screening battery which included the Gates MacGinitie Reading Test (GMRT) of Reading Comprehension, Test of Silent Word Reading Fluency (TOSWRF; Mather, Hammill, Allen, & Roberts, 2004) and a Teacher Reading Rating (TRR; Speece et al., 2010). For the TRR, teachers rated students' reading ability on a scale of 1 to 5; for students rated as 1 or 2 (below grade level), teachers identified the number of problem areas for the student (decoding/word reading, fluency, comprehension, vocabulary, motivation). Raw scores from GMRT Reading Comprehension and TOSWRF and the number of reading problems from the TRR were entered into a logistic regression equation to determine the probability of reading risk. Students with a predicted probability of risk $\geq .40$ were selected to participate. These students were rank ordered by probability and pair matched within school, and then each pair was randomly assigned to condition (i.e., intervention or control).

Intervention description. The students in the treatment group participated in a supplemental Tier 2 reading intervention, which consisted of 24 scripted lessons delivered over 12-15 weeks beginning in January. Tutors met with groups of two to three students 3 days per week for twenty-four 40 minute sessions. Due to attrition, one tutor worked one-on-one with a student, but those lessons were not used for this analysis, due to differences in the interaction style which may have occurred as a function of individual, as opposed to small group, tutoring. Of the 23 groups analyzed as part of this study, 13 of the groups consisted of two students and 10 groups consisted of three students. Differences in how tutors responded which could be related to the size of the group are considered in the results.

During each session, students spent approximately 5-7 minutes on fluency practice by engaging in repeated reading and approximately 30 minutes on reading and understanding expository texts. Tutors used the remaining time to review the day's agenda at the beginning of the lesson and review what was taught at the end of the lesson. Providing cognitive modeling and multiple opportunities for practice, the tutors provided instruction on five comprehension strategies: (a) previewing expository texts, (b) monitoring for understanding, (c) using strategies for decoding unfamiliar words, (d) finding the main idea using paragraph shrinking (Fuchs, Fuchs, Mathes, & Simmons,

1997; Jenkins, Heliotis, Stein, & Haynes, 1987), and (e) question and answer relationships (QAR; Raphael & Au, 2005) (Ritchey, Silverman, Montanaro, Speece & Schatschneider, 2012). Tutors taught each strategy in sets of three or more lessons, in the order they are listed in above. The tutor introduced the new strategy during the first lesson of the set by modeling the strategy and providing opportunities for the students to practice. In the two lessons that followed the introductory lesson, tutors continued to provide modeling and guided practice of the targeted strategy. For example, the lessons examined in this study focused on paragraph shrinking. The paragraph shrinking strategy was introduced in Lesson 13, and Lessons 14 and 15 provided opportunities for additional modeling and guided practice on the paragraph shrinking strategy. A sample lesson is included in Appendix B and the complete scope and sequence of the intervention can be found in Appendix C.

Results. Results from the intervention study indicated that students who received the 16-hour intervention performed better than their control group peers on one measure of reading comprehension— a researcher-created measure, ASKIT, which assessed the ability to apply reading comprehension strategies taught in the intervention (Ritchey et al., 2012). There were no significant differences between the intervention or control

groups on the other measures of reading comprehension, or for measures of reading fluency and word reading.

Current Study

Comprehension instruction was the focus of the present study, since the fluency portion of instruction allowed for very little variation. Since tutors gave feedback during fluency instruction using a standardized script, there was little variation in how tutors responded to students in this section of the lesson. The tutor listened to one student read, marked errors on a copy of the passage, and once the student completed the passage, followed a three step standard protocol for feedback: first telling the student the number of words read correctly, offering a compliment or praise, and identifying one area recommended for improvement. (e.g., “You read 138/142 words correctly today. You read with good phrasing and expression. Next time, be sure to pay attention to the ending letters of words, sometimes you missed endings like –s or –ed because you were trying to read so quickly.”) This highly scripted feedback procedure for responding to student errors during fluency practice left little room for variation. However, during comprehension instruction, although the standardized lesson plan indicated which sections of text to use for modeling and guided practice, as well as provided a list of questions to be asked during the lesson, there was no indication of how tutors should

react to students if they could not independently complete the targeted strategy or answer the scripted questions correctly. When students answered questions incorrectly or displayed a misunderstanding, the tutors needed to make spontaneous instructional decisions while implementing the scripted intervention.

Lesson selection. To control for differences in instructional behaviors associated with the teaching of a specific skill, two lessons focusing on paragraph shrinking – a strategy for identifying main idea – were analyzed. This strategy was selected because tutors indicated in their anecdotal notes in tutor logs that students had difficulty learning this skill; therefore, it was hypothesized that tutors may have had more opportunities to respond to student errors in these lessons.

The paragraph shrinking strategy (Fuchs et al., 1997; Jenkins, Heliotis, Stein, & Haynes, 1987) consisted of four distinct steps: (1) reading the paragraph; (2) identifying the subject of the paragraph (the “who” or the “what” the paragraph is about); (3) identifying the most important thing about the subject; and then (4) saying the main idea in 10 words or fewer. The tutors noted that students had difficulty in both identifying the subject and the most important thing, as well as distinguishing the most important thing from details. A sample lesson including fidelity criteria for each lesson segment is included in Appendix B. The paragraph shrinking strategy was introduced in lesson 13

and practiced again in lessons 14 and 15. In lesson 13, the tutor was directed to explain the strategy and model its use, as well as allow the students an opportunity to apply the strategy. I selected to transcribe and analyze the two lessons that followed the introduction of the strategy, Lessons 14 and 15. These two lessons were selected because they offered more opportunities for student practice, and thus the most potential for variety in student and tutor response.

In Lessons 14 and 15, the students read two expository texts: *Tropical Rainforests* (Sayre, 2003) and *Tropical Rainforest Mammals* (Landau, 1996). Both books were on Guided Reading Level M and Developmental Reading Assessment (DRA) Level 20-24, which roughly equate to texts appropriate for students in second or third grade.

Participants

Tutors. Tutors were 12 graduate research assistants who worked on the larger project over two academic years 2007-2008 and 2008-2009. One tutor was an instructor for both years. Six tutors had prior classroom teaching experience ($M = 4$ years; range = 1-7 years), and three tutors had student teaching experience, though they had never been a teacher of record. Additional demographic information is available in Table 1. I served as one of the tutors in Year 1. The second year, I served as the Intervention Coordinator. In that role, I co-facilitated tutor training, observed and provided feedback to tutors in the

field, and conducted live and audio fidelity checks of tutor adherence to the intervention protocol. A research assistant was hired to code the lessons for which I served as the tutor.

Tutor training. In both years, the tutors participated in approximately 20 hours of training and were required to demonstrate at least 90% adherence to the tutoring protocol prior to the start of the intervention. Training took place over four days, followed by individually scheduled practice sessions. The first day of training included an introductory session in which information about the purpose of the intervention and how students were selected for intervention was shared. One co-investigator led the training both years and communicated expectations of the tutors. Specifically, tutors were to: (a) complete each lesson in 40 minutes; (b) record each session using a digital audio recorder, and (c) include all components listed in the lesson plan. The investigator instructed the tutors to familiarize themselves with the lessons, which were scripted. Tutors did not need to read the script verbatim, but needed to address all components that were listed in the fidelity criteria column of the lesson (see Appendix B for an example). The tutors were given binders with the 24 lesson plans, an intervention scope and sequence, and copies of the texts students would use during repeated reading. Tutors also received copies of the trade books used in the first 12 lessons and were given time to read

the books and familiarize themselves with the content. Tutor training also included explicit instruction and practice of all intervention procedures, including conducting repeated reading individually and with partners, modeling new strategies, and responding to student oral reading errors. Tutors were given time to practice delivering the scripted lessons in pairs, and needed to demonstrate 90% adherence to the tutoring protocol before they were allowed to begin intervention delivery with students.

Training in year 2 was modified based on issues that were identified during year 1 of the intervention. For example, during training, tutors were given a list of ways to responding to student errors (such as repeating the question, or giving prompts). Despite this training, the tutors in year 1 provided significantly more scaffolds ($M = 29.7$) than tutors in year 2 ($M = 22.1$) ($F = 10.88$; $p = .002$). There were no other significant differences between the two groups in terms of experience or other instructional variables.

Table 1
Tutor Demographics

	Tutors (n = 12)	
	n	%
Gender		
Male	4	33.3
Female	8	66.6
Race		
African American	1	8.3
Caucasian	10	83.3
Biracial or Other	1	8.3
Undergraduate Major		
Elementary Education	5	41.6
English	2	16.6
Other (e.g., Business, History, Anthropology)	5	41.6

Students. Students (n = 56) were at-risk readers in parochial schools in a large city and surrounding suburbs in the mid-Atlantic United States who were assigned to receive an intervention. Students were selected for intervention using a screening procedure to identify students who were at risk for reading failure (Speece et al., 2010). Two cohorts of students participated in the intervention. There were no differences by cohort. Demographic information for the students is detailed in Table 2.

Role of Researcher

In qualitative inquiry, the researcher serves as an instrument of the research and interprets the data through their own lens of personal experience, beliefs, and prior

understandings (Creswell, 2007). My role in the research evolved throughout the study from full participant to observer as participant (Glesne, 2006). In year 1 of the larger intervention validation study I served as a tutor, and thus, I was a full participant. This perspective was valuable in that I had a deep understanding of the difficulties I and the other tutors faced during implementation of the intervention. In the second year of the intervention, I was promoted to the role of intervention coordinator. In this role, I assisted in the training, scheduling, and provided instructional support for the tutors. As an experienced teacher myself, I came to this study with beliefs about what constituted good and poor teaching. My beliefs about the intervention were shaped by my own experiences as a participating tutor. Consequently, my interpretations are colored by my previous experiences as a teacher, tutor, and intervention coordinator.

Table 2
Student Demographics

	Intervention Group (n = 56)	
	n	%
Gender		
Male	33	58.9
Female	23	41.1
Race		
African American	25	44.6
Caucasian	25	44.6
Biracial or Other	5	8.9
Mother's Education		
< High School Graduate	1	1.8
High School Graduate	39	70.9
College Degree	10	18.2
Graduate Degree	5	9.1

Note. Some percentages do not add up to 100% due to missing values. The percentages for Race and Mother's Education are based on the number of parents reporting, not the total in the sample.

Design

In order to fully explore the types of off-script instructional behaviors tutors exhibit in response to student oral responses during the implementation of the scripted reading intervention, and the relationship between these behaviors and student outcomes, I employed a mixed methods approach. Mixing qualitative and quantitative approaches provided the best method to fully answer the research questions. Some researchers believe using multiple methods is superior to using a monomethod approach because the strengths of the methods are complementary, and the methodological weaknesses can offset each other (Teddlie & Tashakkori, 2003). It has been well-documented that

scholars disagree which paradigmatic perspectives can, or should, underlie a mixed methods approach (Teddlie & Tashakkori, 2003), or if methods should even be mixed at all (Rossman & Wilson, 1985) although several researchers have argued mixed methods is rooted in pragmatic perspective (Datta, 1997; Rossman & Wilson, 1985; Tashakkori & Teddlie, 1998). The pragmatist view emphasizes practicality, rejects the false dichotomy of needing to subscribe to one paradigm over another (Tashakkori & Teddlie 1998), and supports the use of whichever method can best answer the research questions of interest (Johnson & Onwuegbuzie, 2004). I selected to use mixed methods in this study in order to best answer my research questions, because multiple methods provided the greatest opportunity to fully explore the spontaneous instructional behaviors of tutors and how those tutor differences might relate to student reading comprehension outcomes. Using qualitative methods allowed me to fully describe the types of instructional behaviors present in the lesson transcripts and explain differences across tutors and groups, while quantitative methods provided the opportunity to determine which behaviors were statistically related to each other and how those instructional behaviors might be related to students' scores on reading comprehension measures.

Because the purpose of the study was ultimately to determine the relationship between spontaneous instructional behaviors and student reading comprehension, I

selected to prioritize quantitative methods. However, in order to gather the numerical data necessary to conduct such an analysis, I needed first to determine the types of instructional behaviors the tutors exhibited when delivering the intervention and examine how tutors differed from one and other. Further, simple counts of the frequencies of behaviors could not fully capture the nuance in how tutors reacted to student errors or misunderstandings. Qualitative methods could provide the rich description necessary to fully explore the phenomenon of interest. The mixed methods design served a complementary purpose, as described by Greene (2001), as the qualitative data would illustrate each of the codes and provide information on the quality, and the quantitative data provided a measure of quality of feedback and allowed me to explore the relationship between tutor feedback and student outcomes.

I employed a sequential design because I needed to conduct the qualitative analysis in order to determine which behaviors should be considered in the quantitative analysis (Creswell et al., 2003). Sequential exploratory studies are used to explore a phenomenon and occur in two or more phases (see figure 1).

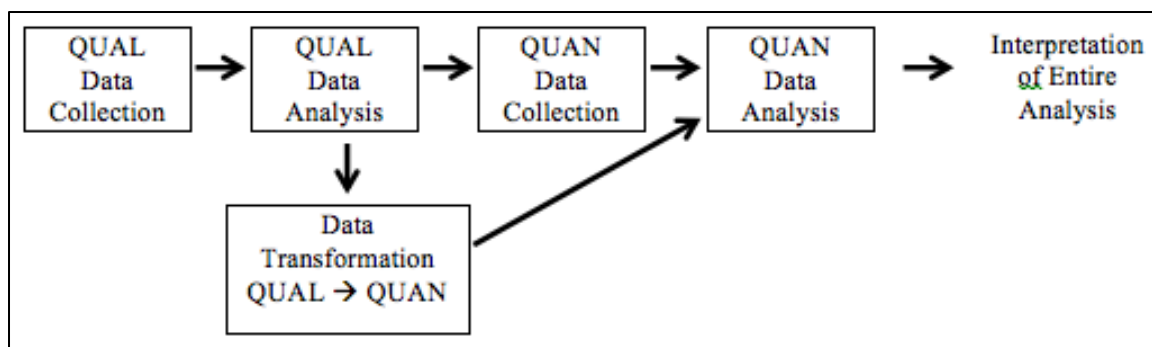


Figure 1. Sequential Exploratory Design (adapted from Creswell et al., 2003, p. 225)

Sequential studies can be exploratory or explanatory. Typically, in sequential explanatory studies, an initial quantitative study is conducted and afterward, qualitative data is collected and analyzed to help explain the quantitative results (QUANT → QUAL). Sequential exploratory designs begin with the use of qualitative methods followed by quantitative methods (QUAL → QUANT). In either design, the researcher can prioritize one method over another (Creswell et al., 2003). In this study, I used a sequential exploratory design in which quantitative methods were prioritized, since the primary question of interest was regarding the relationship between specific instructional behaviors and student reading outcomes. In order to identify which behaviors were important, I needed to begin with a qualitative analysis of the tutor's behaviors. Quantitative data were derived from counting the frequency of specific behaviors.

Qualitative Analysis

For each of the 24 groups, two audio recordings of lessons were transcribed verbatim. Since the study began with an examination of prior research on the relationship between teachers' instructional behaviors and student outcomes, I used directed content analysis (Hsieh & Shannon, 2005) to analyze the data. Potter and Levine-Donnerstein (1999) describe this approach to content analysis as deductive, since existing research was used to design the coding scheme. Initial codes were based on the five behaviors cited in the literature that have demonstrated relationships with student reading achievement, and included: (a) explicitness (Duffy et al., 1986; Hoffman et al., 1984; Lara & Medley, 1987; and Wharton-McDonald et al., 1998); (b) metacognitive modeling (Duffy et al., 1986; Mariage, 1995; and Wharton-McDonald et al., 1998); (c) use of scaffolding (Mariage, 1995; Taylor et al., 2005; Wharton-McDonald et al., 1998); (d) terminal and sustaining feedback (Hoffman et al., 1984; Mariage, 1995; Martin et al., 1980; Taylor et al., 2000) and (e) management (Taylor et al., 2000; Wharton-McDonald et al., 1998) Table 3 displays the five common behaviors identified in the literature.

Table 3
Common Instructional Behaviors Associated with Student Achievement in Reviewed Literature

	Duffy et al. (1986)	Hoffman et al. (1984)	Lara & Medley (1987)	Mariage (1995)	Martin et al. (1980)	Taylor et al. (2000)	Taylor et al. (2005)	Wharton-McDonald et al. (1996)
Explicitness	X	X	X					X
Metacognitive Modeling	X			X				X
Scaffolding				X			X	X
Terminal & Sustaining Feedback		X		X	X	X		
Management						X		X

Coding was not limited to the five behaviors identified in the literature, since not all observed tutor behaviors corresponded to these predetermined codes. Coding was flexible to allow for new codes to describe interactions that did not match these predetermined codes.

In dialogue, each turn by a speaker may be considered an utterance (Levinson, 1983). Each tutor utterance was coded using Process Coding, to indicate the action in the data (Corbin & Strauss, 2008, p. 96-97). Because the utterances of the tutors often included more than one action, Simultaneous Coding (Miles & Huberman, 1994) was employed. Simultaneous coding refers to the use of multiple codes for a single data unit

(Saldana, 2009), in this case, tutor utterance. For example, during one conversational “turn” a tutor may have (a) praised a student’s correct answer, (b) expanded on it by providing additional information, and then (c) asked the next question in the script. In this case, three codes were necessary to capture each of the actions for that single utterance. Use of Simultaneous Coding is appropriate because complex “social interaction does not occur in neat, isolated units” (Glesne, 2006, p.150). Additionally, Simultaneous Coding allowed for the exploration of patterns of behavior within the tutor’s responses to the students.

Throughout coding, I used categorical aggregation (Stake, 1995) to develop categories based on collective instances. After establishing categories, I employed the constant comparative approach (Glaser, 1965; Glaser & Strauss, 1967; Strauss & Corbin, 1990) to evaluate each instance relative to each category to ensure the appropriateness of both the codes and categories. This led to two additional analytic steps: (1) condensing codes for behaviors that seemed topographically similar and (2) splitting codes that could benefit from more in-depth description. For example, all instances of scaffolding were originally coded SCAF; however, while comparing multiple instances it was apparent that the behavior labeled “scaffolding” took on multiple forms. This code was then split

into several subcodes, which provided a better description of the behaviors that made up scaffolding (See Table 7 for the nine sub-coded types of scaffolding observed).

Due to the scripted nature of the intervention, much of the instructional talk was standard across the tutors. Behaviors that were part of the scripted lesson were coded SCRIPT, and were not considered as part of the analysis, since the actions of interest were the "off script" spontaneous instructional behaviors individual tutors introduced while implementing a standardized intervention.

Trustworthiness and Reliability. Several steps were taken to ensure the quality of data collection and analysis. Directed content analysis and peer debriefing are associated with enhancing the credibility (i.e., trustworthiness) of a study (Zhang & Wildemuth, 2009). Coding began with directed content analysis (Hsieh & Shannon, 2005), which built on existing research to develop the coding scheme. Peer debriefing was used during the initial coding process. A doctoral student used the initial codebook to code several lessons, and we met weekly to discuss the themes we discovered in the transcripts. She served as both a peer debriefer and research assistant by “asking hard questions about methods, meanings, and interpretations” (Creswell, 2007, p. 208). We kept notes for each of our peer debriefing sessions.

Implementing consistent study processes and creating an audit trail of memos describing the data analytic procedure can enhance the trustworthiness of a study (Zhang & Wildemuth, 2009). Throughout the development of the codes, my research assistant and I coded the same transcript multiple times. During each iteration, coding disagreements were discussed, definitions and examples refined, and final codes agreed upon. Once initial codes were developed using a small set of lesson transcripts, we applied them to additional lesson transcripts. As more lessons were examined, additional codes were developed to describe interactions that had not been present in previous lessons. Throughout coding, my research assistant and I kept memos about each lesson and highlighted any interactions that were of particular interest. These interactions were discussed during weekly meetings that occurred over 12 weeks. Analytic memos detail what occurred at each meeting.

Once the final codebook was created, all of the lessons were recoded using the finalized coding scheme. My research assistant coded a random sample of 15% ($n = 6$) of the lessons using the finalized list of codes and compared them to lessons I had coded. Intercoder reliability was calculated by dividing the number of agreements by the sum of disagreements and agreements multiplied by 100. Inter-coder reliability was 91%. See Appendix D for the final codebook.

Extreme case analysis. Once frequencies for each of the behaviors of interest were obtained, I was able to examine extreme cases (i.e., the tutor who engaged in the identified behaviors most and least frequently). Caracelli and Greene (1993) describe the extreme case analysis approach as an integrative strategy for mixed-methods research in which cases are identified through analysis of one data type –in this case, quantitative frequencies – and then further examined through data of another type, in this case, qualitative analysis. Extreme case analysis was used to fully address the second research question (RQ2: How do tutors differ from each other in regard to the type, variety, and frequency of the identified spontaneous instructional behaviors?) and develop a more in-depth understanding of the phenomena of interest.

To determine extreme cases, I identified the tutors who exhibited each of the identified spontaneous instructional behaviors with the greatest and least frequency. For each behavior of interest, two tutors were identified for further analysis –the tutor who exhibited that behavior with the most frequency and the tutor who exhibited that behavior with the least frequency. These extreme cases illustrate the variation across tutors in respect to the spontaneous instructional behaviors being studied.

Quantitative Analysis

After answering the first three research questions using qualitative and mixed methods, I used quantitative methodology to answer the remaining research question: How are the spontaneous instructional behaviors of tutors related to student reading comprehension outcomes? The qualitative analysis yielded four distinct spontaneous instructional behaviors, all of which typically occurred after a student was unable to correctly respond to a question posed by the tutor: (a) scaffolding (i.e., using questions or prompts to lead the student to the answer); (b) simply telling the student the correct answer; (c) providing unclear feedback about the correctness of the answer; and (d) erroneous feedback. Frequency counts were used to transform the qualitative codes into numerical data, a process Tashakkori and Teddlie (1998) refer to as “quantitizing” (p. 126). Each independent variable that emerged from the qualitative analysis is described below. The results section includes a more detailed description and illustrative examples for each of the spontaneous instructional behaviors identified. Frequency, as opposed to rate, was deemed appropriate because of consistency across tutors in the length of time spent in comprehension instruction, which ranged from 30.5 to 33 minutes.

Independent variables. I included four independent variables were included in the qualitative analysis.

Frequency of scaffolding. The number of instances of scaffolding identified through the qualitative analysis was determined by summing the number of questions or prompts that were intended to assist the student in correctly exhibiting the comprehension skill being taught or to answer a question correctly. Nine subtypes of scaffolding were identified through qualitative coding of the transcripts: (a) giving sentence starters; (b) use of simple directives; (c) asking leading questions; (d) breaking the task into smaller steps, then prompting the student to complete the steps; (e) providing multiple choice-type options; (f) reminding the student of content previously taught; (g) making connections to contexts familiar to the students, (h) repeating or rephrasing the question or task directions; and (i) asking metacognitive questions. Instances of these nine scaffolding behaviors were summed to yield a frequency count of scaffolding behaviors.

Frequency of telling. Occurrences where the tutor simply provided the answer to the students after posing a question or demand were totaled to yield a count of the number of instances of telling.

Frequency of unclear feedback. Unclear feedback was defined as tutor feedback that did not clearly communicate whether the answer given by the student was correct or incorrect. I counted the total number of episodes of unclear feedback in each lesson.

Frequency of erroneous feedback. The frequency of erroneous feedback was characterized by three responses to students: (a) responding to an incorrect answer as if it were correct, (b) ignoring an error all together, or (c) responding to the student's correct answer as if it were an error.

Dependent variables. Three student outcome measures of reading comprehension were administered. Two measures (GMRT and Maze) were administered prior to the intervention (pretest) and after the intervention (posttest), while the other (ASKIT) was administered at posttest only.

Gates-MacGinitie Reading Test, Fourth Edition (GMRT). The GMRT Reading Comprehension subtest (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) was administered in a group setting. The GMRT is a timed assessment; students have 35 minutes to silently read short narrative and expository passages and answer multiple-choice questions. Form S was administered for pre-test, and Form T was administered for post-test. The alternate form reliability and the test-retest reliability correlations exceed .90 for fourth grade students. Concurrent validity for GMRT has been established at .75 for grade 4 with Maze (Jenkins & Jewell, 1993) and .73 with the Metropolitan Achievement Test 6th edition (MAT-6; Jenkins & Jewell, 1993).

Maze. Maze (Fuchs, n.d.) was group administered at pre-test and post-test as an

assessment of silent reading and comprehension. The task uses a modified cloze technique. Every seventh word following the first sentence of a reading passage is deleted and replaced with three choices. Students select the word appropriate to the context of the passage. In two minutes, students must complete as many sentences as possible. Both reliability and criterion validity of the Maze are strong ($r = .60$ to $.86$; Fuchs & Fuchs, 1992). Test–retest reliability has been reported at $.83$ (Manchester et al., 2004). Maze scores have been demonstrated to have strong correlations ($r = .77$) with the Reading Comprehension subtest of the Stanford Achievement Test (Fuchs & Fuchs, 1990 as cited in Wayma, Wallace, Wiley, Tichá, & Espin, 2007). Students completed two grade 4 probes. Scores from the two probes were averaged, and then transformed to reflect the mean of the number of correct words selected per minute.

Assessment of Strategy Use and Knowledge for Information Text (ASKIT). The Assessment of Strategy Knowledge and Use for Information Text (ASKIT; Ritchey, Speece, Silverman, & Montanaro, n.d.) was designed specifically for this intervention to assess students' knowledge of comprehension strategies and ability to apply these strategies while reading expository texts. It was individually administered only at post-test. Students answer questions about reading strategies (previewing, identifying the main idea, retelling, summarizing) and demonstrate these strategies while reading an

authentic information text. The expository text used (*Protecting Sea Turtles*, O'Sullivan, 2004) differed in content from those used in the intervention. Knowledge and strategy questions are scored using a 0 to 3 scale, with each score defined with criterion and examples. Appendix E includes sample items and examples of corresponding scored responses. Two raters scored the items with inter-rater agreement exceeding .90. Any scoring discrepancies were discussed and a final score for that item agreed upon.

Cronbach's alpha for this sample was .56 ($n = 123$), just below the recommended .60 for researcher-developed measures (Gersten et al., 2005). Reliability for this measure is acceptable for research purposes given the restricted range with this at-risk sample. The ASKIT is devised of two scores: percentage of words read correctly (untimed Reading Accuracy) and total raw score (Comprehension). Only the comprehension score was used in this study, since the relationship between instructional variables and student reading comprehension scores was the topic of interest. The comprehension portion of the assessment requires that the students explain the strategies they use to preview texts and figure out unfamiliar words, retell, summarize, and find the main idea of passages in an expository text. A copy of the measure and scoring examples are included in Appendix E.

Control variables. The following variables were used as control variables: (1) pretest score (for the two measures administered pre and posttest –Gates MacGinitie and

Maze); (2) parents' level of education; (3) racial background; and fidelity of implementation.

Parents' level of education. Multiple studies have established the association between parental level of education and their child's level of academic achievement (Bakker, Denessen, & Brus-Laeven, 2007; Spera, 2006; Schlechter & Milevsky, 2010; U.S. Department of Education, 2000). In fact, Myrberg and Rosen (2008) found that parental level of education accounted for as much as 58% of the variance in student reading achievement. On a student background questionnaire, parents of participants were asked to identify the parents' level of education as (1) High school graduate or less; (2) Some college/vocational training; (3) College degree; or (4) Graduate degree.

Race. Multiple studies have revealed that differences in academic achievement can be attributed to race (Bali & Alvarez, 2004; Brooks-Gunn, Klebanov, & Duncan, 1996; Willie, 2001). Although a better explanation for this achievement gap may actually be related to factors related to school quality, test bias, or cultural differences among several other factors, some studies have found that these achievement differences remain even after controlling for home and school factors (Fryer & Levitt, 2002). In order to control for any differences attributable to racial or ethnic differences, it was included as a fixed factor in the regression models.

Parents were asked to supply race information on the student background questionnaire. Parents were asked to identify their child's race or ethnicity as (a) American Indian, (b) Alaskan Native, (c) Asian, (d) Black or African American, (e) Native Hawaiian, (f) Pacific Islander, (g) White, (h) Hispanic or Latino or (i) more than one race. In the final data set, race or ethnicity was collapsed into three categories: Black, White, and biracial or other.

Fidelity of implementation. Adherence to the scripted protocol for comprehension instruction was calculated by counting the number of components present and dividing by the number of possible components. In order to determine this, I analyzed the lesson scripts and identified the key components of the lesson (i.e., the tutor modeling the strategy, providing opportunities for guided practice, or discussing key vocabulary). The data sheets used to assess fidelity of implementation for the comprehension portion of two lessons of interest are included in Appendix F.

Analysis. Frequency counts of the observed instructional behaviors were entered into a regression equation predicting student growth controlling for student demographic factors and prior achievement. Ordinary least squares (OLS) regression with robust standard errors (Cohen, Cohen, West, & Aiken, 2003) was used in this study to examine the relationship between quality of feedback and student outcomes. This method was

selected in order to account for students clustered within tutor. Although the structure of the data is nested, the sample size of the existing data set precluded the use of hierarchical linear modeling (HLM). Each of the 12 tutors worked with groups of 2-4 students pulled from 15 classrooms from 12 schools. Due to scheduling, some tutors taught several small groups, while others taught only one. Tutors worked with few students ($M = 4.75$ per tutor; range = 2-8) rendering multilevel modeling inappropriate, given that the suggested benchmark is a minimum of 30 groups with 30 individuals per group (Kreft, 1996).

Power analysis. Given the fixed number of participants, I conducted a post-hoc power analysis using G-Power to reduce the probability of a Type II error. With a sample of 56 students, an alpha level of 0.05, and nine predictors, achieved power to detect an effect size of 0.15 was calculated to be 0.8101, which is considered adequate power (Huck, 2008).

Assumptions. Prior to analysis, I examined the data to ensure it met the assumptions of linearity and normality for linear regression. To assess the presence of outliers I generated the ratio between the skewness and kurtosis coefficients with their respective standard errors, and compared them to the acceptable ranges, suggested by Kendall and Stuart (1958). The skewness-standard error ratio should be within +/- 2

while kurtosis should be ± 5 (Kendall & Stuart, 1958). To assess multicollinearity I examined the correlation matrix and VIF (variance inflation factor). Bivariate correlations greater than .90 and VIF values greater than 5 indicate multicollinearity (O'Brien, 2007). The results of the preliminary exploratory analysis of the data are detailed in Chapter IV.

Statistical tests. Hierarchical multiple regression with robust standard errors was selected to examine the relationship between tutor behaviors and student outcomes. The initial model included control variables, including race and parents' level of education. The independent variables were the following instructional behaviors: (a) the number of feedback errors made by tutors, (b) the number of scaffolding prompts, (c) the frequency of non-instructional talk, (d) the number of instances of unclear feedback, and (e) the number of answers that were simply told to the student.

Regression Models. What follows are the regression models used to predict the outcome variables, including relevant null and alternative hypotheses, as well as descriptions of the predictors included at each level of modeling.

Regression models for predicting GMRT and Maze. The following models were used for the two outcome variables that had pretest scores – GMRT Comprehension and Maze.

$$\text{Reduced Model (1a): } \hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i}$$

$$\text{Model 1b: } \hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i}$$

$$\text{Model 1c: } \hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i}$$

$$\text{Model 1d: } \hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i}$$

$$\text{Model 1e: } \hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i}$$

Full Model: (Model 1f)

$$\hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i}$$

Hypotheses:

$$H_0: \Delta P^2 = 0$$

$$H_A: \Delta P^2 \neq 0$$

Table 4
Predictors for GMRT Comprehension and Maze

Reduced Model Predictors (Model 1a)
X ₁ = Race
X ₂ = Parents' level of education
X ₃ = Pretest Score
Model 1b
X ₁ = Race
X ₂ = Parents' level of education
X ₃ = Pretest Score
X ₄ = Fidelity
Model 1c
X ₁ = Race
X ₂ = Parents' level of education
X ₃ = Pretest Score
X ₄ = Fidelity
X ₅ = Scaffolds
Model 1d
X ₁ = Race
X ₂ = Parents' level of education
X ₃ = Pretest Score
X ₄ = Fidelity
X ₅ = Scaffolds
X ₆ = Tutor Tell
Model 1e
X ₁ = Race
X ₂ = Parents' level of education
X ₃ = Pretest Score
X ₄ = Fidelity
X ₅ = Scaffolds
X ₆ = Tutor Tell
X ₇ = Unclear Feedback
Full Model Predictors (Model 1f)
X ₁ = Race
X ₂ = Parents' level of education
X ₃ = Pretest Score
X ₄ = Fidelity
X ₅ = Scaffolds
X ₆ = Tutor Tell
X ₇ = Unclear Feedback
X ₈ = Erroneous Feedback

Regression models for predicting scores for ASKIT. The following models were
be used for the ASKIT, the dependent variable for which there are only posttest scores.

$$\text{Reduced Model (1a): } \hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}$$

$$\text{Model 1b: } \hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i}$$

$$\text{Model 1c: } \hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i}$$

$$\text{Model 1d: } \hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i}$$

$$\text{Model 1e: } \hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i}$$

Full Model: (Model 1f)

$$\hat{Y}_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i}$$

Hypotheses:

$$H_0: \Delta P^2 = 0$$

$$H_A: \Delta P^2 \neq 0$$

Table 5 Predictors for ASKIT	
Reduced Model Predictors (Model 1a)	
X ₁	= Race
X ₂	= Parents' level of education
Model 1b	
X ₁	= Race
X ₂	= Parents' level of education
X ₃	= Fidelity
Model 1c	
X ₁	= Race
X ₂	= Parents' level of education
X ₃	= Fidelity
X ₄	= Non-Instructional Talk
Model 1d	
X ₁	= Race
X ₂	= Parents' level of education
X ₃	= Fidelity
X ₄	= Scaffolds
Model 1e	
X ₁	= Race
X ₂	= Parents' level of education
X ₃	= Fidelity
X ₄	= Scaffolds
X ₅	= Tutor Tell
Model 1f	
X ₁	= Race
X ₂	= Parents' level of education
X ₃	= Fidelity
X ₄	= Scaffolds
X ₅	= Tutor Tell
X ₆	= Unclear Feedback
Full Model Predictors (Model 1g)	
X ₁	= Race
X ₂	= Parents' level of education
X ₃	= Fidelity
X ₄	= Scaffolds
X ₅	= Tutor Tell
X ₆	= Unclear Feedback
X ₇	= Erroneous Feedback

Summary

Transcripts of reading comprehension lessons were qualitatively coded for spontaneous tutor behaviors to determine the types of various instructional behaviors tutors used in reaction to student oral responses during a scripted reading intervention. Qualitative differences in how the tutors interacted with students during instruction were analyzed. Tutors' adherence to the standard protocol was also explored by determining

the percentage of lesson components performed in the session. Frequency counts of the instructional behaviors identified through qualitative analysis were entered into statistical models predicting students' reading comprehension scores to determine the relationship between student reading comprehension and the spontaneous instructional behaviors exhibited by the tutors who provided instruction, after controlling for student-level demographic characteristics and prior achievement.

Chapter IV: Results

This study explored the types of spontaneous instructional behaviors tutors exhibit in reaction to student oral responses during a scripted reading intervention. Data included verbatim transcripts of two lessons for 23 groups of fourth grade struggling readers taught by 12 tutors. The tutors were all trained graduate assistants, some of whom had previous teaching experience. The two lessons selected for analysis focused on the comprehension strategy paragraph shrinking, in which students were taught to identify main idea of a paragraph in 10 words or fewer by identifying the topic (i.e., the who or what the paragraph was about) and the most important thing about the topic. Tutors indicated on tutor logs that their students struggled to learn the paragraph shrinking strategy, which made these two lessons ideal for exploring how tutors spontaneously reacted to student errors since the scripted lessons provided little guidance in how to respond to student difficulties.

I used descriptive qualitative codes to identify the types of spontaneous instructional behaviors the tutors used in reaction to student oral responses. To determine the frequency of each of the identified behaviors, I tallied and summed each occurrence of the qualitative codes. To explore differences across tutors I used extreme case analysis, identifying cases using quantitative frequencies (i.e., tutors who displayed behaviors the

most or least frequently) and descriptively analyzing those extreme cases. In order to determine the relationship between the identified behaviors and student outcomes, frequency counts for each behavior were entered into multiple regression equations predicting student outcomes for three measures of reading comprehension administered at post test. Results are presented in this chapter, which is organized by research question:

1. What types of spontaneous instructional behaviors do tutors exhibit in reaction to student oral responses within a scripted reading comprehension intervention?
2. How do tutors differ from each other in regard to the type, variety, and frequency of the identified spontaneous instructional behaviors?
3. How do tutors differ from each other in terms of their fidelity of implementation of the scripted lesson plan?
4. What is the relationship between the spontaneous instructional behaviors of tutors and student reading comprehension outcomes, controlling for student demographic variables, prior achievement, and fidelity of implementation?

Please note that all student and tutor names throughout this manuscript are pseudonyms.

Research Question 1

The first research question involved identifying the types of spontaneous instructional behaviors that tutors exhibited in reaction to student oral responses while implementing the scripted reading intervention. Descriptive qualitative coding was used to identify how tutors responded to students during instruction. Coding suggested that tutors typically followed the script closely until the student made an error or had a misunderstanding. The script provided no guidance for the tutor if the student had difficulty “shrinking” the paragraph to express the main idea in 10 or fewer words, or if they failed to answer a question posed by the tutor correctly. Tutors responded to these student errors in four main ways: (a) scaffolding, which included using questions or prompts to lead the student to the answer; (b) telling the student the correct answer; (c) providing unclear feedback about the correctness of the answer; or (d) providing erroneous feedback. The number of occurrences for each code was tallied. Mean frequencies for each behavior is presented in Table 6.

Scaffolding was the behavior the tutors used with the greatest frequency, followed by telling students the answer. Providing unclear feedback or erroneous feedback occurred far less often. The following sections detail each of the instructional behaviors tutors exhibited in reaction to students’ oral responses during the scripted lessons. The

behaviors are presented in order of descending frequency and are supported by descriptive examples. In each of the examples provided, the tutors' pseudonym is bolded, correct answers are presented following tutors' questions in parenthesis, and student errors are noted.

Table 6
Frequency of Instructional Behaviors

Instructional Behavior	Mean Frequency (SD)
Scaffolding	26.34 (10.33)
Tell	5.38 (3.75)
Unclear Feedback	3.35 (2.59)
Erroneous Feedback	1.8 (1.83)

Scaffolding. Most frequently, tutors responded to students by scaffolding.

Scaffolding was defined as questions, prompts, or instructional talk intended to guide a student toward a correct answer. Nine types of scaffolding emerged from the data, as detailed in Table 7.

Table 7
Types of Scaffolding

Type	Definition	Examples
1. Starter	Tutor provides the first words of sentence to prompt the student to complete the sentence.	“Orangutans have...” “So tropical rainforests have...”
2. Directive	Tutor tells the student to do something specific which will assist them in answering the question posed.	“Look at the picture to help you” “Look at your chart.”
3. Leading	Tutor asks a question, or series of questions, or prompts in an effort to lead the student toward the correct response.	“Where is that happening?” “What’s the other kind?”
4. Steps	Tutor asked a question or gave a prompt that broke down a process into steps or stages. Frequently used when students were implementing a strategy that included several steps (such as paragraph shrinking).	“What do you do next?” “So what’s the next step?”
5. Multiple-Choice	Tutor posed a question followed by offering two or more choices.	“Is it hot or is it cold?” “In different places or in the same places?”
6. Reminder	Tutor provided a reminder of a past lesson, something said earlier in the lesson, or about something that appeared in the text.	“Remember how we talked about the different layers of the forest?”
7. Connection	Tutor asked a question or gave a prompt, which required the student to make connection between something in their life and the concept presented.	“If your mom says fix your appearance before you head out to church, what is she trying to say? Fix what?”
8. Rephrase	Rephrase the original question.	[What is a characteristic or a feature of a mammal?] “What makes a mammal a mammal?”
9. Self-Evaluate	Tutor posed a “metacognitive” type question encouraging the student to evaluate answer	“Is that really what’s most important?”

The two most common types of scaffolding tutors used were starters and steps. Tutors often gave students prompts that started the response they were looking for the student to complete. In this example a tutor uses a starter when reviewing the steps of the paragraph shrinking strategy.

AMANDA: Paragraph shrinking. So what am I doing first? First I— (SCAF-starter)

SEAN: Read.

The tutor started the sentence that they wanted the student to finish.

Similarly, this tutor offers a starter to help her students draw comparisons about the two kinds of forests they read about:

ROBIN: So basically this paragraph is kind of comparing two different kinds of forests. What do they two forests have in common? They're both— (SCAF-starter)

CRAIG: They're both tropical forests.

Starters provided a support for students to respond to the tutors' questions in that it prompted the student to complete the statement begun by the tutor.

Another type of scaffold that was common was prompting students to use the steps of the strategies they had been taught. Prompts encouraged students to approach reading in a strategic way by following set procedures for interacting with texts. For example, when students previewed a text, the tutor asked, "So now that we've looked at the text features what should we do next in previewing?" This type of prompt encouraged the student to vocalize the strategies they were using during previewing. Another example of prompting for using the step-by-step procedures is illustrated in this interaction:

ROBIN: Right. So Craig just told us the most important thing about the who or the what. And then what's the last step, Mike?

MIKE: Say the main idea and tell it in ten words.

In this example, Robin, the tutor, had the students label each steps of the strategy as they practiced applying it in text. This type of scaffold, prompting for steps, encouraged the

students to explicitly rehearse the names of the steps as they practiced the paragraph shrinking strategy. Each of the tutors used this type of scaffold in the lessons.

Metacognitive scaffolding occurred least frequently. Metacognitive scaffolds encouraged the students to evaluate their own responses. Only three tutors used metacognitive scaffolds (Marnie, Robin, and Doug), and it is noted that each of those tutors had previous professional teaching experience. In this example illustrating a tutor's use of metacognitive scaffolding, the tutor, Robin, asks Craig to identify the most important thing about rainforests.

CRAIG: The most important thing about the tropical rainforest is that the huge amount of rainfall.

ROBIN: Is that the most important, or is it a detail? (SCAF-META)

CHRIS: It's a detail.

Chris did not identify the most important thing about the rainforest according to the paragraph he just read, but Robin uses a question to encourage Chris to self-evaluate the correctness of his answer, rather than correcting him herself.

Other tutors used similar types of questions to encourage self-evaluation, asking students, "Is that the main idea of the paragraph?" or "Is there anything you would like to add that you think would be important?" The few tutors who did use this type of self-evaluative scaffolding used it sparingly, no more than one or two times during a lesson. Self-evaluative scaffolds could be useful to help students acquire a self-regulatory internal

voice that they could use to evaluate the completeness or correctness of their own responses.

Tutor tells the answer. Another type of behavior observed in reaction to students' oral responses were coded "Tutor Tell". When students were not able to provide a correct answer, sometimes the tutor would simply tell them the correct answer. At times, this occurred after a failed attempt or multiple attempts to scaffold students toward the correct answer. In this first example, the tutor gave the student several opportunities to answer the question, and used numerous scaffolds before eventually telling the student answer.

- MARNIE:** So first step is to read, go ahead. Start to think. So now we're gonna stop and name the person, animal, or thing the text is about. (ANSWER: Scientists)
- CARLY:** In Peru, this is Peru. (ERROR)
- MARNIE:** Peru is a place mentioned in one sentence. Okay, so who's the person in this?
- CARLY:** Scientists.
- MARNIE:** Scientist is the person, so we'll think okay, that's the who. So what's the most important thing that this says, this whole big chunk of the paragraph? What about scientists? What are they doing? (ANSWER: Find many species of plants and animals in the rain forest)
- CARLY:** They found 43 different ant species.
- MARNIE:** That they're finding a lot of different species, right? It says, like that last sentence says, "Every year scientists find new plants and animals in the rainforest." So we have our "who or what"-- the person, animal, or thing --The who is scientists. And what's the most important thing? What are they always doing?
- CARLY:** They're always looking.
- MARNIE:** Looking for what?
- CARLY:** Insects and stuff?

- MARNIE:** Insects and animals and plants. Okay, so do you think we could do ten words or less about that?
- CARLY:** Yes.
- MARNIE:** Okay.
- CARLY:** Ant species— (ERROR)
- MARNIE:** You gotta start with the person, animal, or thing it's about.
- CARLY:** Okay. Ants can be seen— (ERROR)
- MARNIE:** Well, our who or what was scientists because this is about scientists. Scientists –
- CARLY:** Scientists spotted 43 species in one tree. (ERROR)
- MARNIE:** Okay. But is that the main idea of *all* the sentences here?
- CARLY:** No.
- MARNIE:** We want to have the main idea of all the sentences.
- CARLY:** The scientist—
- MARNIE:** Scientists—
- CARLY:** --scientists in Peru found 43—43 different species in one tree. (ERROR)
- MARNIE:** Okay, but that's only—that's telling me one sentence. See how you're still only telling me one sentence?
- CARLY:** Scientists found 43 different –
- MARNIE:** Let's not stick with the 43, let's stick with just—kick out the 43. It's a detail.
- CARLY:** Scientists found different ant species on one single tree but that tree—no—(ERROR)
- MARNIE:** See how we're going? It's hard to say just the main idea, without the details --with all the information. So what did we scientists do?
- CARLY:** A scientist—
- MARNIE:** Mya told me, what did they find?
- CARLY:** Scientists found different plants—
- MARNIE:** Different plants. Different animals, different—
- CARLY:** Ants.
- MARNIE:** --different ants, different animals, different plants. So can we say that? Scientists—

CARLY: Scientists found 43—

MARNIE: We don't want to go into details. Because we want it to be short.

CARLY: Scientists in Peru—(ERROR)

MARNIE: No, we don't want details. Peru is a detail.

CARLY: Scientists –

MARNIE: Found—

CARLY: -found different ant species—(ERROR) --found a cont—a species bigger than...(ERROR)

MARNIE: What did they do? Scientists found different plants, animals, and insects, right, in where? (TELL)

CARLY: In—

MARNIE: The—

CARLY: --in the

MARNIE: Rain—

CARLY: --rainforest.

After multiple unsuccessful attempts, the tutor finally provides the answer and even attempts to scaffold the last portion, perhaps in an effort to have the student feel successful, despite needing to be told the answer. Most tutors did not allow so many opportunities before telling the student the correct answer. Providing so many scaffolded opportunities may not have been helpful since this interaction wasted instructional time that might have been better utilized by tutor modeling of the strategy. The time spent interacting with this one student could have also led to the other student in the group to lose interest.

Here, the tutor gives each of the children in the group an opportunity to answer before providing the correct answer.

- DOUG:** So what is the person, animal, or thing that this is about? Darren?
(ANSWER: Rainforest)
- DARREN:** Tiny monkeys called marmosets? (ERROR)
- DOUG:** Is that what this—is that what this paragraph is about?
- DARREN:** No.
- DOUG:** Kimberly, what do you think it's about? What do you think this paragraph is about?
- KIMBERLY:** It's about, hold on, let me get it first. Monkeys and marmosets. Mammalsets. (ERROR)
- DOUG:** We're not talking about the picture with the caption, we're talking about this paragraph that I just read. Simon, what do you think?
- SIMON:** It's about the animals, they look for insects, ants, and plants, and birds? (ERROR)
- DOUG:** Oh, oh, no, that's not quite right.
- SIMON:** The rainforest settle down in the—
- DOUG:** It's about the rainforest is what it's about. It's about the rainforest. (TELL)

The tutor attempted to have the students answer the question, but once he had given each student an opportunity, he supplied the answer himself. This offered the tutor an opportunity to assess if one or more students in the group understood how to identify the topic of the paragraph. The students' responses indicated that none of the students understood how to identify the topic correctly.

In some instances, tutors provided the correct answer immediately following an error. Typically, this occurred when students were asked questions in which they were offered two discrete choices, such as indicating if a mammal is warm or cold blooded.

Once the student answered incorrectly, the answer was obviously the one they had not selected. This type of immediate telling also occurred when students were asked to classify an animals' adaptation as either a physical or behavioral adaptation.

SAMUEL: Are these physical or behavioral adaptations?

ERIC: Behavioral. (ERROR)

SAMUEL: Physical. (TELL)

Simply telling the student the correct answer did not allow the student to understand why their initial response was incorrect. Other tutors reacted to similar incorrect responses by rewording the question “Are these physical or behavioral adaptations?” asking instead “Does it have to do with how the animal looks or how the animal behaves?” Often when the question was reworded in such a way student were able to correctly answer the question, indicating that they may not yet have been comfortable with the vocabulary *behavioral* or *physical*, but understood the difference between the two concepts.

Unclear Feedback. Unclear feedback was defined as feedback delivered to students that did not clearly communicate if the answer given was correct or incorrect. At times, tutors seemed hesitant to communicate to the students that their answers were incorrect. Unclear feedback was often followed by another type of behavior – telling the student the answer, calling on another student, or using scaffolding. Interestingly, in

several examples of unclear feedback the tutor uses the phrase “kind of” in response to the student’s incorrect answer. Unclear feedback could be confusing to students who are struggling to master new skills, such as identifying the main idea using paragraph shrinking. When students are learning new skills, they rely on feedback to gauge if they are progressing successfully. If that feedback is not clear or does not provide an honest assessment of their success, students may have a difficult time adjusting their performance to increase their success.

For example, here a tutor asks the student to identify the “who or what” for the paragraph they just read, but when the student answers incorrectly the tutor seems hesitant to directly communicate that the answer is not correct.

- JENNIFER:** Do you guys think you can pick out what is the main—the most important—or what is the paragraph talking about? I’ll give you a chance. (ANSWER: The rainforest)
- RYAN:** Tiny monkey? (ERROR)
- JENNIFER:** Tiny monkey, well, kind of. (UNCLEAR) What do you think it’s talking about, Liam?
- RYAN:** Or like different kinds of animals?
- LIAM:** What most animals do in the rainforest.

The feedback to Ryan –“kind of”– was unclear because it did not indicate what was problematic about the answer provided, or what was “kind of” correct about the answer.

The tutor communicated that the answer was incorrect indirectly by calling on another student to respond, a type of terminal feedback.

The next example of unclear feedback is similar in that the tutor, Abbie, uses the same phrase “kind of” in response to the student’s incorrect answer; however in this example, the tutor uses another type of terminal feedback, a tutor tell, to communicate the correct answer.

- ABBIE:** What is the main idea? So put that in ten words of your own words. (ANSWER: Rainforests have many types of plants)
- SERENA: There are many plants and animals in different rainforests. (ERROR)
- ABBIE:** Kind of, almost, yeah. (UNCLEAR) So you could just say rainforests have many different kinds of plants. (TELL)
- SERENA: And animals.
- ABBIE:** Well, it doesn’t really talk about animals in this paragraph.

At the end of the interaction the tutor clarifies the reason the student’s original answer was “almost” correct. It seems the tutor was reluctant to directly address the faulty part of the student’s answer, until the student insisted on keeping animals as part of the main idea. Only then did the tutor communicate the reason the answer was incorrect.

In this final example of unclear feedback, the tutor follows the unclear feedback with a multiple-choice type scaffold.

- DOUG:** Great, scientists, good job. And what is the main idea? Who knows what the main idea is, David?
- DAVID: About studying animals. (ERROR)
- DOUG:** About studying animals? So that’s more or less... (UNCLEAR)
- DAVID: About species.
- DOUG:** Yeah, about what kind of—a lot of species or a few species? (SCAF-MC)

Erroneous feedback. The spontaneous instructional behavior that occurred with the least frequency was providing erroneous feedback. Some tutors gave incorrect feedback in response to student answers. Erroneous feedback included three types of tutor errors: (a) praising incorrect answers, (b) ignoring errors, or (c) treating the correct answer as if it were incorrect.

Praising incorrect answers. The first way tutors provided erroneous feedback was to praise an incorrect answer as if it was correct. Here, a tutor asks a student to define a vocabulary word, *suited*, which had been introduced in that lesson.

KATHERINE: Its body is well suited for this way of life. And we just said that well suited means--?

ANNA: Comfortable. (ERROR)

KATHERINE: Perfect, good, suitable, very good. (ERRONEOUS FEEDBACK)
[Begins reading from text] Its curved fingers grasp tree branches easily. Its long arms are ideal for swinging through the forest.

When the student provided an incorrect synonym, *comfortable* for *suited* the tutor praised the student, communicated that the answer was ‘perfect’. She did not provide a correct definition for *suited* or draw the student’s attention to the correct synonym, *ideal*, which appeared in the text they were reading. Thus, the student could leave the lesson with the faulty understanding that *suited* is a synonym for comfortable.

Several of the tutors exhibited this behavior of praising incorrect answers. Here, another tutor discusses the importance of pollination.

MICHAEL: And so again, these animals are the ones that go to one plant, they take the pollen from that plant and put it to the next one, and to the next one, and to the next one, okay? Why is it important to the rainforest that this [pollination] happens?

MARTIN: So it can protect others. (ERROR)

MICHAEL: Well, kind of, yes, Martin, that's very good. Very, very good. (ERRONEOUS FEEDBACK) But without pollinators plants won't get the right nutrients. They need—the other plants need that pollen, right. And so they need it to spread it around so that everyone can get energy so that they can bear, you know, so a lot of those trees can put, bear fruit so that we can eat, you know?

This tutor praises the incorrect answer by saying “That’s very good. Very, very good,” although it is not clear how pollination is related to “protecting others” or what exactly the “it” was that Martin was referring to. The tutor then goes on to provide a factually incorrect answer, referring to the pollen as nutrients the plants need to survive, rather than explaining the role pollen plays in plant reproduction. As a result of this interaction, the student will continue to misunderstand the role of pollen in sustaining plant life, and leave with the understanding that pollen provides some sort of protection or food for the plant.

In the next example, the tutor praises a student’s successive approximation of using the paragraph shrinking strategy, in which the students are to identify the main idea of a passage in 10 words or fewer. The tutor praises the student for using nine words in her summarization; however, the tutor does not make it clear that the student did not

correctly identify the main idea of the passage, which described the two seasons that occur in seasonal rainforests.

KATHERINE: What is it telling us—okay, I know, what is it telling us about seasonal rainforests?

ERICA: Seasonal tropical forests have many kinds of different trees.
(ERROR)

KATHERINE: Excellent, that's nine words, very good, Erica, very, very good.
(ERRONEOUS FEEDBACK) Or you could also say that seasonal rainforests have a dry season and a wet season. (TELL)

The student gave an incorrect answer, yet the tutor praised her answer. The tutor did not definitively communicate that the answer is incorrect, but couched her feedback by saying, “you could also say” followed by telling her the correct answer. There are two possible explanations for why tutors praised incorrect answers –the tutors may not have known how to identify the main idea of a passage well enough themselves. Another possibility is that perhaps since the students were identified as struggling readers, the tutors did not want to discourage the students by pointing out their errors.

Ignoring errors. Another way tutors provided erroneous feedback was to ignore the response given by the student altogether. Here the tutor, Doug, asked Lara to paragraph shrink. Lara identified the most important thing, but the tutor suggests she include additional information. Then the student asked a question, which the tutor ignores.

DOUG: Okay. What's the most important thing about the orangutan in this?

- LARA:** They—people have to kill the mother to get the baby.
(PARTIALLY CORRECT)
- DOUG:** Okay. So baby orangutan's mothers are killed. Anything else? Are they becoming, like are the numbers shrinking, too, right?
- LARA:** Um.
- DOUG:** Okay, let's tell that in ten words or less.
- LARA:** The last one. Do people have to kill the mother for the baby?
- DOUG:** Okay. George, you want to continue reading, please, buddy?
(IGNORE)

The tutor does not respond appropriately to Lara's question, and the tutor does not address the fact that Lara never finished shrinking the paragraph as directed. The tutor ignores Lara's response and asks the other student in the group to begin orally reading the next section of text. This demonstrates something unique to small group, as opposed to one-on-one tutoring. In a small group setting, the tutors needed to attend to multiple students simultaneously and maintain lesson momentum. Further, the timed nature of the lessons (i.e., needing to adhere strictly to the 40 minute session and complete all activities on the script) may have caused the tutor to move on to the next portion of the text before ensuring the understanding of each of the students.

Here, a tutor demonstrates a pattern of ignoring both the responses of the students. This tutor seemed to stick to the script so closely, he barely responded to the students at all, which is illustrated in this interaction.

- SAMUEL:** Okay, guys. We can save it till the end if we have time. But

we only have so much time. Are these physical or behavioral adaptations? (ANSWER: PHYSICAL)

- ELI: Behavioral. (ERROR)
- SAMUEL: Physical. (TELL)
- ELI: Wait, how is it physical because—
- CHASE: Because his body parts—We’re talking about his claws and stuff.
- ELI: Oh, oh, ‘cause I thought you were talking about the grasp.
- CHASE: No. They don’t live on grass, you idiot.
- SAMUEL: So its curved fingers grasp the tree branches easily [*reading text*]. And its long arms—
- CHASE: Why does it have curved fingers? What does it look like? Does it look—
- SAMUEL: (interrupts, reading) His long arms are ideal for swinging through the forest. Okay.
- ELI: It’s like this hand, or like...?
- SAMUEL: So how is your body well suited for your way of life?

In the above interaction, Eli indicated he did not understand why his answer *behavioral* was incorrect and Chase asked several questions about the orangutan, yet the tutor ignored both and asked the next question on the script (How is your body well suited for your way of life?). One possible reason for the tutor moving on so quickly, despite students’ misunderstandings or questions, is that the tutors were trained to complete all of the lesson components in 40 minutes.

Treating the correct answer as if it were incorrect. The final way tutors gave erroneous feedback was in response to a student’s correct answer, treating it as if it were wrong. This behavior occurred much less frequently than treating an incorrect answer as

if it were correct. In the following example, the tutor responds as if Maya's answer is incorrect, although her answer is correct. The text she had just read detailed how orangutans are often called *red heads* because the orangutan has long shaggy reddish or orange hair.

KATHERINE: What makes an orangutan look different from other monkeys?

ANYA: Their hair.

KATHERINE: Okay. Thinking about the information that you've read, the information is right there, you don't need to guess.

Summary. In reaction to students' oral responses, tutors exhibited four main types of spontaneous instructional behaviors: (a) scaffolding –using questions or prompts to lead the student to the answer; (b) telling the student the correct answer; (c) providing unclear feedback about the correctness of the answer; or (d) providing erroneous feedback. Sometimes these behaviors occurred together within the same interaction. Scaffolding occurred most frequently, and nine subtypes of scaffolding were identified. Tutors sometimes provided the answers to the students, either immediately following an incorrect answer or after several failed attempts at scaffolding. Less often, tutors seemed hesitant to directly communicate to students that their errors were incorrect, and provided either unclear feedback or erroneous feedback, providing praise for incorrect answers.

Infrequently, tutors provided erroneous feedback in which they treated a student's correct response as if it were incorrect.

Research Question 2

The second research question was: How do tutors differ from each other in regard to type, variety, and frequency of the identified spontaneous instructional behaviors? This question was answered using a combination of quantitative and qualitative data. For each behavior, I used quantitative frequencies to identify extreme cases. The extreme cases selected for each behavior were the tutors who exhibited that behavior with the greatest frequency and least frequency. For each of the cases identified, I present description and illustration of interactions typical for that tutor. When considering these cases, it is important to keep in mind that these cases are not representative of all tutors, but are rather a means of describing the variation between the tutors in terms of the frequency with which they exhibited the spontaneous instructional behaviors of interest.

Tutors differed in the frequency with which they demonstrated the behaviors identified in research question 1: (a) scaffolding, (b) telling, (c) providing unclear feedback, and (d) giving erroneous feedback. Table 8 details the frequencies for each tutor/group combination.

Scaffolding. Tutors differed in terms of the frequency of scaffolding they used, which seemed to depend on the group of students with whom they were working. Across all lessons and tutors, the mean frequency of scaffolds was 26.35 (SD=10.33), and ranged from 6.5 – 47.

Tutors also differed in the variety of types of scaffolding they used. Some tutors tended to use only three or four types of scaffolds again and again, while others seemed to have a larger repertoire of scaffolding behaviors. For example, Robin, a certified teacher and reading specialist-in-training with 3 years professional teaching experience, seemed to skillfully select from all nine types of scaffolding as appropriate, while Abbie, a graduate assistant with no previous teaching experience, demonstrated only three types of scaffolding: (a) sentence starters, (b) prompting for steps, and (c) rewording the question.

Table 8
Mean Frequency of Behaviors Across Lessons

Tutor	Scaffolding	Tell Answer	Unclear Feedback	Erroneous Feedback
Katherine				
Group 1	34	8	6	2
Group 2	16	4.5	2.5	3.5
Group 3	24	15	6	2
Logan				
Group 4	22	2	0	3
Marnie				
Group 5	19	2	0	3
Group 6	31.5	2.5	0.5	0.5
Group 7	17	2.5	0	1.5
Doug				
Group 8	37	2.5	8	3.5
Group 9	25	9	3	0
Group 10	33	7	7	8
Jennifer				
Group 11	12	7	3	1.5
Amanda				
Group 12	47	1	5	1
Group 13	40	3	6	0
Group 14	25.5	1	2	0.5
Robin				
Group 15	27.5	1	0.5	0
Melissa				
Group 16	25	2	0	0
Group 17	13.5	7.5	4	0.5
Stacey				
Group 18	13	6	2	1
Michael				
Group 19	27.5	5	2	2
Group 20	17	5	0.5	2
Samuel				
Group 21	19.5	11	5	5
Abbie				
Group 22	21	9	3.5	1.5
Group 23	6.5	5	4.5	0

In fact, by examining Table 9, it is clear that teaching experience does seem to be related to the number of types of scaffolding tutors have in their repertoire. More

experienced tutors demonstrated more variety in the types of scaffolds they used when interacting with their students. In fact, there was a strong significant correlation ($r = .869$) between the number of the types of scaffolding demonstrated and number of years of professional teaching experience. Years of experience also had a moderate positive correlation to the number of scaffolding opportunities which were taken advantage of, which was computed by taking the number of scaffolding prompts demonstrated by the tutor and dividing by the number of student errors made in the lesson.

Table 9
Tutor Experience and Scaffolding Repertoire

Tutor	Years of Professional Teaching Experience	Student Teaching Experience? (Y/N)	Types of Scaffolding Demonstrated
Amanda	7	Y	9
Marnie	6	Y	9
Katherine	5	N	9
Robin	3	Y	9
Doug	2	Y	7
Melissa	1	Y	6
Logan	0	N	5
Jennifer	0	Y	5
Stacey	0	Y	5
Michael	0	Y	3
Samuel	0	N	5
Abbie	0	N	3

Frequent versus infrequent scaffolders. Tutors differed in the frequency with which they used scaffolding in response to students' oral responses. By examining the mean number of scaffolds by each tutor across the two lessons, I identified tutors who were frequent scaffolders (FS) versus those who were infrequent scaffolders (IS). Robin

was a frequent scaffolder ($M=27.5$) who rarely told students the answer ($M=1$). In response to a student error, Robin provided feedback that identified why the answer the student had given was incorrect followed by an explanation. She seemed able to pinpoint exactly what student misunderstanding needed to be addressed. For example, in response to a student who was listing details, rather than the main idea, Robin responded: “Okay. I think those would be included, maybe. Those are specific. When we do a main idea we want to be more general. So what is this whole thing about? It’s about the—.” In this exchange, Robin acknowledges the student’s response, but explains that he has provided details, rather than the more general main idea. After providing the student with an explanation of why his answer was not correct, she follows up with two scaffolds – a leading question (“So what is this whole thing about?”) and a sentence starter (“It’s about the—”).

Robin’s use of scaffolds was quite skillful. For example, when a student gave an answer that did not match what was said in the text, she directed the student to look back in the text. While other tutors often gave directives which seemed focused on the error (i.e., “Is that what the book said? Read it again.”), Robin tended to use more gentle directives focused on collaboratively verifying meaning with the text (i.e., “I might have missed that. Let’s look in the text and see if that’s right.”).

Robin used a variety of scaffolds while teaching. She demonstrated all nine types of scaffolds in each of the lessons she taught. One type of scaffolding Robin used which was unusual among the tutors in this study was self-evaluative questioning. This type of questioning modeled for students how to ask themselves the type of questions needed to monitor the correctness of their own answers. For example, after a student gave an answer, Robin would ask: "Is that the most important, or is it a detail?" Her students were able to self-assess the correctness of their own responses through her use of self-evaluative questions. She even used self-evaluative questioning preemptively, that is, before a student provided an incorrect answer. Sometimes before calling on students she would ask them to self-assess the quality of their answer asking questions such as "Are you sure you're right?" or "So you're ready to say it in 10 words or less?"

One reason Robin may have had such a wide variety of scaffolding strategies and applied each of them so skillfully was because she had several years of teaching experience. The other frequent scaffolders were also experienced teachers (Katherine, 5 years experience, $M = 24.7$; Marnie, 6 years experience, $M = 22.5$). In contrast, two of the most infrequent scaffolders, Jennifer ($M = 12$) and Stacey ($M = 13$) had no professional teaching experience. In a post hoc analysis of the relationship between years

of professional teaching experience and variety in types of scaffolding demonstrated, the two were found to be strongly correlated ($r = .869$).

When examining the infrequent scaffolders, it became evident that these tutors tended to do more telling. When students gave an incorrect answer, these infrequent scaffolders tended to simply tell the student the answer, thus terminating the interaction, rather than provide scaffolding and sustaining the interaction. This pattern was not consistent for all tutors, but was evident for tutors who fell on the extreme ends of the scaffolding spectrum.

Both infrequent scaffolders primarily used two types of scaffolds: leading questions and prompted use of steps. Both Jennifer and Stacey used three other types of scaffolding infrequently: sentence starters, directives, and reminders. In comparison to Robin, they seemed to have a limited repertoire of scaffolding behaviors to choose from when responding to a student's incorrect answer.

Another difference between the frequent scaffolders (FS; Robin, Katherine, and Marnie) and infrequent scaffolders (IF; Jennifer and Stacey) was the persistence with which they used scaffolding. The frequent scaffolders often used multiple types of scaffolds in succession in an effort to have the student eventually answer the question, as illustrated in this interaction between Robin (FS):

- ROBIN:** I don't even know that the most important who or what is even written here. Oh, yes, it is, I see it. (Answer: Rainforest)
- STUDENT:** Ants, plants, and birds. (ERROR)
- ROBIN:** Okay. I think those would be included, maybe. Those are specific. When we do a main idea we want to be more general. So what is this whole thing about? It's about the— (*SCAF-Leading Question; SCAF-Sentence Starter*)
- STUDENT:** Animals.
- ROBIN:** It's about the animals where? (*SCAF-Leading Question*)
- STUDENT:** In the tropical rainforest.
- ROBIN:** Okay. So maybe we could say that this is about the— (*SCAF-Sentence Starter*)
- STUDENT:** The rainforest?
- ROBIN:** Yeah, it's about the rainforest. So we could use our erasers to circle "rainforest." Rain—look, buddy, rainforest. Okay. And you kind of already told me what the main idea is about the rainforest. I've noticed that sometimes you skip this stuff and you just go right to the main idea. You don't necessarily know who or what it's about. But you already started telling me, what about the rainforest? There's all kinds of— (*SCAF-Leading Question; SCAF- Sentence Starter*)
- STUDENT:** Animals.
- ROBIN:** And—you said ants, what are ants? Ants are— (*SCAF-Leading Question; SCAF- Sentence Starter*)
- STUDENT:** Insects.
- ROBIN:** Insects and— (*SCAF- Sentence Starter*)
- STUDENT:** People.
- ROBIN:** And people. So how could we say that together? (*SCAF-META*)
- STUDENT:** Ants, insects—
- ROBIN:** Let's start with the first step, the most important thing. What was the most important thing? (*SCAF- Steps*)
- STUDENT:** The rainforest has ants and—
- ROBIN:** Or we could just say insects. (*SCAF-Leading*)
- STUDENT:** I mean insects—
- ROBIN:** And then we get ants and butterflies taken care of. Insects— *SCAF-Sentence Starter*)
- STUDENT:** Insects, people, and animals.
- ROBIN:** Perfect, seven words. That's awesome. That's exactly what we want to say.

Robin is persistent in her attempts to have the student answer correctly, and provides multiple scaffolds until the student is able to produce the main idea with her guidance. In contrast, consider these interactions, which are typical of the interactions between the infrequent scaffolders and their students:

- JENNIFER:** So what do you think is the most important thing that this paragraph is about? (ANSWER: Rainforest)
- STUDENT:** About birds. (ERROR)
- JENNIFER:** Maybe. I don't think---that was the first thing. This one's really tricky, I'm gonna go ahead and tell you. The thing that this paragraph is about is the rainforest, right? (TELL)
- STUDENT:** Uh-huh. (Yes).

Jennifer asks the student to answer, the student offers an incorrect response, and instead of scaffolding, Jennifer tells the student the answer.

Stacey, another infrequent scaffolder responded similarly to student errors, by simply telling the student the answer following an error.

- STACEY:** What are we talking about? Usually it says it in the first paragraph? (ANSWER: Scientists)
- STUDENT:** Animals.
- STACEY:** No, we're talking about scientists in there, too.

Students made errors that provided their tutors an opportunity to use scaffolding; however these opportunities were often missed. In this next example, instead of telling the student the answer, Stacey responds to Heather's error by making noises similar to those you might hear on a game show:

- STACEY:** Okay, so we did step one. Step two, what's the person, animal, or thing that this paragraph is about? What are we talking about in here? Heather? (Answer: Scientists)
- HEATHER:** Plants and animals.
- STACEY:** [Makes noise like a buzzer]
- HEATHER:** The rainforest
- STACEY:** [Makes noise like a buzzer]
- CASSIE:** Scientists.
- STACEY:** Yeah, you got it, ding ding ding.

Stacey “buzzes” Heather’s incorrect answer and subsequent incorrect guess, until another student in the group, Cassie, provides the correct answer. The interaction does not move Heather toward understanding why her answer and subsequent guess were not correct, or provide her with an opportunity to correct her answer.

Tutor Tell. Tutors also differed in how often they told students the answer.

Across all groups and tutors, the mean frequency of tutor tells was 5.38 instances per lesson (SD = 3.75). All tutors told students an answer at least once in each lesson. The average number of tutor tells ranged from 1-11 tells per lesson. Some tutors were consistent in their frequency of telling students the answer across all of the groups they taught, such as Marnie (2, 2.5, and 2.5) and Michael (5, 5), while other tutors seemed to use different rates of telling depending on the group of students they were teaching. For example, Doug seemed to use a higher rate of telling with two of his groups (Group 9, nine tells; Group 10, seven tells), but with his third he used telling less often (M=2.5).

Frequent versus infrequent tellers. To determine which tutors tended to tell more or less frequently than others, I averaged the number of times tutors used telling across groups. One tutor, Samuel, told students the answer more frequently than others. On average, Samuel told students the correct answer 11 times per lesson, as compared to the rest of the tutors ($M=5.38$). In contrast, some tutors, such as Robin ($M=1.0$), Logan ($M=2.0$), and Marnie ($M=2.3$) tended to rarely tell students the answer.

While Samuel was the tutor who most frequently supplied the correct answer for the students, he was also the tutor who most frequently supplied the wrong answer or gave unclear feedback. One reason Samuel may have been the most frequent teller is because he gave less than the average amount of scaffolding. Instead of using scaffolding when a student gave an incorrect answer, Samuel tended to just tell his students the answer.

Unclear feedback. Unclear feedback occurred less often than the other spontaneous instructional behaviors ($M = 3.35$, $SD = 2.59$). Unclear feedback ranged from 0 to 8 occurrences per lesson. The amount of unclear feedback seemed to be relatively stable within tutor, as the frequencies did not seem to differ very much across the different groups. That is, it seems the tendency to give unclear feedback may be

characteristic of the tutor's communication style, rather than a function of the interaction between tutor and students.

Clear versus unclear feedback. In examining the frequency of unclear feedback, I identified tutors who were more and less clear with their feedback. The tutors with the fewest number of instances of unclear feedback were Logan (M=0), Marnie (M=0.17), and Robin (M=0.5). The tutors with the highest levels of unclear feedback were Samuel (M=5) and Katherine (M=4.83). Tutors who were clear with their feedback tended to communicate to the student not only that their answer was incorrect, but *why* it was incorrect, as in this example:

ROBIN: Okay. So based on that, if you want to reread it you can. What do you think the most important thing is in that paragraph?

STUDENT: That nuts fall—

ROBIN: Hold on a second, buddy, you're telling me the main idea [most important thing].

STUDENT: Oh.

ROBIN: I want you to think of the who or the what first, okay?

In contrast, tutors who gave unclear feedback tended to ask unclear questions or make comments that were less direct. For example, here Samuel asks the students in his group to tell the main idea of the paragraph they just read:

SAMUEL: Okay. Chase, that was step number one, we just read it. Okay. Step number two, stop and name the person, animal, or thing the text is about. What is this text about? This is for you, Chase.

CHASE: This is about orangutans.

SAMUEL: Very good. What's the most—

CHASE: And how they live.

- SAMUEL:** Good. What's the most important thing that this paragraph is telling us? What's the main idea?
- CHASE:** The main idea about this paragraph is about the orangutan.
- SAMUEL:** Okay. But what's the main idea? (UNCLEAR QUESTION)
- ELI:** But what about it?
- CHASE:** The main idea is—
- SAMUEL:** I mean, I'm trying to figure this out, too.
- CHASE:** --is that they don't go in groups.
- SAMUEL:** Okay, that could be one of the main, main ideas. Is there— (UNCLEAR)
- ELI:** Now, say something in ten words or less the most important thing.
- SAMUEL:** Is there another—well—

Samuel seems uncertain of the main idea himself, as he admits he too is trying to figure it out. The other student in the group, Eli, seems to have taken on the role of coach by offering more direct scaffolds to the other student in the group. The tutor, Samuel, repeats the question, “But what’s the main idea?” and Eli clarifies, “But what about it?” One potential reason tutors may have given unclear feedback is because of their own lack of confidence with the material.

Another reason for unclear feedback could be ambiguity in the tutor’s language when teaching or practicing the strategy. Paragraph shrinking is a strategy for teaching students how to find the main idea. A paragraph shrink has two parts (1) the who or what and (2) the most important thing. In this example, Katherine asks Sean to shrink the paragraph they just read. Notice how she uses the term “main idea” to refer to both parts of the paragraph shrink.

KATHERINE: What's the main idea? What is he talking about?

SEAN: The rainforest.

KATHERINE: Okay, good. And what's the main idea?

SEAN: Shrink it now?

KATHERINE: Yeah, I want you to shrink it. What's the main idea?

SEAN: That the rainforest –

KATHERINE: First tell me the main idea and then we'll try and shrink it.

SEAN: The main idea is -- main idea is about the rainforest.

KATHERINE: What about it?

SEAN: It's one kind of a tropical rainforest. This, it has green areas, it's a huge amount of rainfall.

KATHERINE: But you're reading word for word. Remember you're trying to get the overall idea. The overall idea. So you gave us -- Sean has given us the main -- the -- what the paragraph is about. It's about the rainforest, the tropical rainforest. Can you give us the main idea that the paragraph is giving us? Try and pay attention, honey.

Katherine started by asking, “What’s the main idea?” and accepted Sean’s answer (the rainforest) and offered praise (“Okay, good”). However, she repeats the question, “And what’s the main idea?” even though she accepted Sean’s answer. Sean seemed uncertain what his tutor was asking for, since his answer to “What’s the main idea?” was already accepted and praised, and asked for clarification on what he should do (“Shrink it now?”). Katherine’s repeated ambiguous use of the term “main idea” did not support Sean in practicing the paragraph shrinking strategy.

Katherine tended to use imprecise language when explaining strategies as well. The strategy the students had previously worked on was retelling, so the scripted lesson asked the tutors to draw a distinction between the previously taught strategy, retelling,

and the new strategy, paragraph shrinking. In this example, Katherine offers her explanation:

KATHERINE: Remember you're retelling this. So if you start giving us all those little details, you're saying exactly what's in the book. So you're trying to say it in your own words. So it's kind of like retelling, but also paragraph shrinking. Want to try it one more time?

First, she tells the student that they are retelling, and so they need to say it in their own words and then explains: "So it's kind of like retelling, but also paragraph shrinking".

From this explanation, it is unclear what the difference is between retelling and paragraph shrinking and what exactly she would like the student to do when they "try it one more time". Compare Katherine's explanation to that of Robin's: "Paragraph shrinking is like retelling, but we only pull out the very most important information or the main idea" or Marnie's explanation: "Paragraph shrinking is just like retelling, except we're only thinking about the main idea so we're not giving all the details." Both Robin and Marnie were able to communicate the difference between retelling and paragraph shrinking.

Erroneous feedback. Erroneous feedback was also a behavior less frequently exhibited by the tutors overall ($M=1.83$, $SD=1.83$). Erroneous feedback ranged from 0 to 8 episodes per lesson. All tutors were observed giving erroneous feedback at least once. Two tutors –Doug and Samuel – gave erroneous feedback at higher rates than the other tutors, while others, such as Robin, Amanda, and Abbie rarely gave erroneous feedback.

As previously described, tutors used erroneous feedback in three ways: (a) praising an incorrect answer, (b) correcting a correct answer, or (c) ignoring a student error altogether. Samuel, who delivered the most erroneous feedback, frequently praised incorrect answers as if they were correct. His feedback to students regarding paragraph shrinking indicated that Samuel was not able to consistently identify the main idea of a paragraph. The primary way Doug provided erroneous feedback was by praising incorrect answers as well. Doug tended to tell students their answers were “pretty good” and move on to the next example. Unlike Samuel, Doug may have recognized the students’ errors, because at one point after telling a student that his incorrect answer was “pretty good,” he said, “We can work on it a little bit more, but that’s pretty good.” Perhaps Doug was aware that the students were struggling to identify the main idea, but was unsure how or even *if* he should change instruction, given the scripted nature of the intervention. Being told that their incorrect answer was “pretty good” could be confusing for students, particularly if students were working on the same strategies in their classrooms. It could be confusing to receive positive feedback from a tutor regarding their progress identifying the main idea and receive critical feedback from their classroom teacher when working on the same comprehension skill.

Another trend that was noted in regard to erroneous feedback was that some tutors praised students for using 10 or fewer words when they gave the main idea, even if the main idea was not correct. In the following example, Doug asks a student to give the main idea of a paragraph (Main idea: The tropical rainforest contains a variety of plant species). The student responds:

STUDENT 1: A forest—wait. A tropical rainforest has long vines and it is colorful.

STUDENT 2: That's 11.

STUDENT 1: Nah-ah, it's ten.

DOUG: Well, just get rid of the "A". Tropical rainforest –

Rather than correct the student's main idea selection, Doug scaffolded the student toward telling the incorrect main idea in 10 or fewer words. At no point did he indicate that the main idea itself was not correct. The tutor's overemphasis on the number of words in the main idea summary focused the students' attention toward restating the incorrect main idea in fewer words, but did not help the student understand how to correctly identify what the most important information of the paragraph. The target of stating the main idea in 10 or fewer words is intended to assist the student in paring down the main idea to just the most essential information presented in the paragraph, but more important than the number of words in the answer is being able to identify the most important information in the paragraph.

Group Size. Another variable that may have influenced tutors responded differently to students during instruction is the size of the group of children being taught. As stated earlier, instruction was delivered in groups of 2 or 3 students. In order to determine if how the tutors responded to the students may have been a function of the size of the group, I conducted a post hoc analysis of the frequencies of behaviors by group size using one-way analyses of variance (ANOVA). There were no differences by group in the frequency of scaffolding ($p = .164$), frequency of unclear feedback ($p = .764$) or frequency of erroneous feedback ($p = .932$). However, the group size did have a significant relationship with the number of times that tutors told the students the answer ($p = .021$). That is, tutors told students the answers more frequently when teaching groups of three students ($M=6.45$) than when teaching groups of two students ($M=4.15$).

Summary. The 12 tutors differed in the frequency with which they exhibited the four tutor behaviors identified –scaffolding, telling, unclear feedback, and erroneous feedback. In terms of scaffolding, tutors who had previous teaching experience tended to use scaffolding more frequently and demonstrated more variety in terms of the types of scaffolding they used while interacting with students. In terms of telling, tutors who had a more limited repertoire of scaffolding behaviors tended to rely on telling the answer to students, rather than providing scaffolding to allow students to answer themselves. In

terms of unclear feedback, some tutors provided clear feedback by directly communicating to students that their answer was incorrect and explaining why the answer was not appropriate; while other tutors seemed reluctant to address student errors, and instead gave students ambiguous feedback. In terms of erroneous feedback, a few tutors who had higher rates of erroneous feedback were most likely to provide praise for incorrect answers, perhaps because of the tutors' lack of mastery with the skills they were teaching.

Research Question 3

Tutors also differed in regard to fidelity of implementation, that is, how closely they followed the script and accomplished each of the essential components of the scripted lesson. Using the fidelity checklists designed for this purpose, I was able to calculate the percentage of components tutors completed in each of the scripted lessons. The checklists are included in Appendix F. Fidelity ranged from 61% to 100%, with a mean of 76.3% (SD = 9.12). On average, tutors completed 76.3% of the essential lesson components in the comprehension section of the lesson. As illustrated in Table 10, some tutors differed in fidelity percentages across the groups they taught, while others were more consistent. For example, Marnie completed 100% of the components for one group,

89% with another group, and 78% with another group, while Melissa and Michael completed the same proportion of components with each of the groups they taught.

Table 10
Fidelity of Implementation Summary

Tutor	Percentage of Components Present
Katherine	
Group 1	78%
Group 2	67%
Group 3	78%
Logan	
Group 4	61%
Marnie	
Group 5	89%
Group 6	78%
Group 7	100%
Doug	
Group 8	67%
Group 9	78%
Group 10	67%
Jennifer	
Group 11	78%
Amanda	
Group 12	78%
Group 13	67%
Group 14	78%
Robin	
Group 15	78%
Melissa	
Group 16	67%
Group 17	67%
Stacey	
Group 18	89%
Michael	
Group 19	78%
Group 20	78%
Samuel	
Group 21	78%
Abbie	
Group 22	89%
Group 23	67%

A closer examination of the number of tutors who completed each of the various components revealed some interesting patterns in terms of which components of the lessons tutors frequently omitted. The three components most frequently overlooked by tutors – modeling, discussing the purpose of the strategy, and providing opportunities for practice– are described in detail below.

Modeling. Each scripted lesson included a segment in which the tutor was supposed to model the strategy for the students. In examining the fidelity components that were most often omitted by tutors, the modeling component was absent in 83% of the lessons. Although each lesson provided an opportunity for the tutor to demonstrate the strategy being taught, in practice, the majority of tutors seemed to complete the example provided for modeling using a guided practice approach, in which the tutor prompted the students to complete the example with assistance. The modeling was designed to provide the students with a demonstration of a skilled reader using the strategy, and included scripted think alouds, which provided a “behind the scenes” view of the metacognitive process used while using the strategy. The scripted lessons included opportunities for the students to engage in guided practice only after they had observed the targeted skill demonstrated by their tutor. In only 17% of the lessons tutors provided explicit modeling of the paragraph shrinking strategy using think alouds, despite having a detailed script for

modeling included in the lesson plan. It is important to note that the strategy was introduced in Lesson 13, which was not analyzed as part of this study. Perhaps the tutors felt that the modeling provided in Lesson 13 was sufficient, and deemed the modeling in Lessons 14 and 15 as unnecessary. Since Lesson 13 was not analyzed, I cannot determine the efficacy with which the tutors modeled the strategy upon introduction. And without tutor interviews, I cannot determine why they may have opted to use a guided practice approach, rather than the explicit modeling that was included in the lesson script.

Here is a sample of the modeling portion of the scripted lesson. Italics denoted what the tutor was to say.

Tutor reads text and models identifying the main idea
pp. 44 (Hand out books). *Turn to page 44.*

Step 1: Read

Read aloud: Meanwhile, somewhere in a rainforest, hummingbirds hover. Pig-like tapirs walk the ground. Butterflies flutter. Tiny monkeys called marmosets snooze. Scientists look and wonder at ants, and plants, and birds. And rainforest people settle down to sleep in their homes.

It sounds like a lot of information, but we can shrink this.

Step 2: What is the person, animal or thing this is about? [The Rainforest]

Step 3: What is the main idea? [lots of different things live in the rainforest, many plants and animals live in the rainforest]

Step 4: Shrink it! Let's say it in 10 words or less. [The rainforest has animals, insects, plants, and people living together]

Paragraph shrinking is like retelling, but we only pull out the most important information or the main idea.

The scripted lesson indicated that the tutor was supposed to model the strategy for the students. The answer for each step of the modeled strategy was provided.

Additionally, the lesson provided a think aloud for the tutor to demonstrate positive self-talk. This type of positive self-talk is frequently incorporated in strategy instruction and is

designed to boost students' self-efficacy in using the strategy (*"It sounds like a lot of information, but we can shrink this"*; see Graham & Harris, 2009). After modeling the strategy, the lesson plan indicated that the tutor should draw a distinction between a strategy the students had encountered in previous lessons, retelling, and paragraph shrinking, the current strategy.

Here is an example of how a tutor presenting the portion of the lesson described in lesson segment above, using a guided-practice approach, rather than modeling the strategy herself:

- KATHERINE:** Okay. So I want to read this aloud, okay? "Meanwhile, somewhere in a rainforest hummingbirds hover, pig like tapirs walk the ground. Butterflies flutter. Tiny monkeys called marmosets snooze. Scientists look and wonder at ants and plants and birds, and rainforest people settle down to sleep in their homes." That sounds like a lot of information, but we can shrink this. Okay, first of all what is person, animal, or thing that this is about? What is this about? (ANSWER: The rainforest)
- STUDENT 1: It's about -- it's about animals. (ERROR)
- KATHERINE:** Where? (SCAF-LEADING Q)
- STUDENT 1: In the rain -- in the rainforest.
- KATHERINE:** So it's about the rainforest, very good. What is the main idea about the rainforest? (SCAF-STEPS)
- STUDENT 1: The animals are -- (ERROR)
- KATHERINE:** Try and find what -- yeah, that this -- (SCAF-DIRECT, interrupted)
- STUDENT 1: There is many animals in the rainforest?
- STUDENT 2: There are monkeys.
- KATHERINE:** That there are many different animals that live in the rainforest. Only --?
- STUDENT 1: And plants.
- KATHERINE:** And plants, right. So let's try to shrink that and say it in less -- in ten or less words. (SCAF-DIRECT)
- STUDENT 1: Oh, okay. There's -- -there's many kinds of animals in the rainforest.
- KATHERINE:** Okay, that's eight words, good. "There's many types of animals in the rainforest." Seven words, actually, but we said it's not only many animals but also plants. So, [STUDENT 2], I want you to try to shrink that paragraph.
- STUDENT 2: Um, there are lots of animals in the rainforest.

KATHERINE: Okay. But you forget the plants as well. (TELL)

STUDENT 2: And the plants.

KATHERINE: And the plants.

STUDENT 2: That's nine.

KATHERINE: Do you want to try it one more time? (SCAF-REPHRASE)

STUDENT 1: Yeah.

KATHERINE: Okay.

STUDENT 1: There are many plants and animals in the rainforest.

KATHERINE: Excellent. You did that in nine words, very good. Good job, guys.

The tutor, Katherine, completed this portion of the lesson by having the students to perform each step of the strategy supported by tutor scaffolding. This section of the lesson, however, was designed for the tutor to model the strategy. Modeling was intended to be an opportunity for the students to observe a more skilled reader (the tutor) perform the strategy; however, the majority of tutors elected to forgo modeling and instead had the students participate in figuring out the main idea of the example intended for modeling.

Providing the reason for the strategy. Another component of the scripted lesson that was often skipped by the tutor was providing a reason for learning the strategy. In the lesson plan, the tutors were directed to ask the students: "How does this [paragraph shrinking] help us be better readers?" However, a discussion of the reason for using paragraph shrinking was present in only 43% of the lesson transcripts.

Providing opportunities for practice. The scripted lesson also directed the tutors to give each student at least one opportunity to practice the paragraph shrinking strategy during the 40 minute lesson, however, in approximately half of the lesson transcripts (48%) not every student was given an opportunity to practice the strategy himself or herself.

Research Question 4

I used hierarchical multiple regression to answer my final research question (RQ4) and determine the relationship between the spontaneous instructional behaviors of tutors and student reading comprehension outcomes. The analysis began with an examination of the data set to assess the underlying assumptions of multiple regression, including the normality and multicollinearity of the data.

Preliminary analysis. The preliminary analysis initially involved generating the descriptive statistics for each of the predictor and outcome variables. The means, standard deviations, and the relevant skewness and kurtosis statistics are presented in Table 11. The absolute value of the skewness statistics for each variable was compared to twice the standard error, a common rule of thumb suggested by Seltman (2012). By this standard, most variables met the assumption of normality, although the distribution for erroneous feedback was moderately positively skewed. However, the skewness and kurtosis

statistics for all variables were within the acceptable range (± 2 ; Huck, 2008) making it unnecessary to perform any transformations before conducting the analysis.

Multicollinearity was assessed by examining the bivariate correlations and VIF (variance inflation factor) of the predictor variables. Table 12 presents the intercorrelations of all the study variables while Table 13 summarizes the VIF values for the predictor variables.

By examining the correlation matrix, I was able to determine the relationship between the predictor and outcome variables. When interpreting correlations in education research, correlation coefficients from between 0.0 to 0.3 can be considered “weak”, coefficients greater than 0.3, but less than 0.6 can be considered “moderate”, and coefficients greater than 0.6 can be considered “strong” (Connolly, 2007, p. 95). The correlation matrix indicated that there were no significant correlations between the predictors and the outcome variables. None of the outcome variables had strong correlations with each other. Bivariate correlations between the predictor variables ranged from zero to moderate. Unclear feedback was weakly, but significantly, correlated with fidelity ($r = -.275$) and significant, moderate correlations with scaffolding ($r = .315$), tutor telling ($r = .445$), and erroneous feedback ($r = .427$). Tutor telling had a weak negative correlation with GMRT posttest scores ($r = -.273$), which was significant.

To further check for the presence of multicollinearity, I examined the tolerance and VIF statistics for each of the full models. VIF statistics are summarized in Table 13. From the table, it should be noted that across all possible combinations, VIF values are within the acceptable range (<3 ; O'Brien, 2007) and all tolerance values exceed .01 (Meyers, Gamst, & Guarino, 2006). Thus, we can be confident that multicollinearity is not a cause for concern.

Table 11
Descriptive Statistics

Variables	M	SD	Skewness		Kurtosis	
			Statistic	SE	Statistic	SE
Predictors:						
Fidelity	76.30	9.12	.484	0.32	.117	0.63
Erroneous Feedback	1.83	1.83	1.55	0.32	2.99	0.63
Told Answer	5.38	3.75	.88	0.32	.29	0.63
Scaffolds	26.35	10.33	.39	0.32	-.13	0.63
Unclear Feedback	3.35	2.60	.28	0.32	-1.09	0.63
Outcomes:						
Gates MacGinitie Score	21.16	7.02	0.77	0.32	0.38	0.63
Maze Score	9.09	4.00	0.56	0.32	-0.29	0.63
ASKIT	24.16	5.02	-0.16	0.32	-0.43	0.63

Table 12
Correlation Matrix

	1	2	3	4	5	6	7	8	9	10
1. GMRT Pre	1.00									
2. GMRT Post	.289	1.00								
3. Maze Pre	.138	.134	1.00							
4. Maze Post	.157	.134	.388**	1.00						
5. ASK IT	.102	.181	.090	.112	1.00					
6. Fidelity of Implementation	.035	.162	.058	-.144	-.175	1.00				
7. Scaffolding	-.068	-.029	-.190	.210	-.074	-.112	1.0			
8. Tutor Tell	-.184	-.273*	-.147	-.035	-.146	.156	-.070	1.0		
9. Unclear Feedback	-.018	-.261	-.136	.168	-.082	-.275*	.315*	.445**	1.0	
10. Erroneous Feedback	-.009	-.171	-.040	.184	.001	-.157	.245	.245	.427**	1.0

Note. *p<.05, ** p<.01

Table 13
Summary of VIF Values

	Tolerance	VIF
GMRT Comprehension		
Fidelity	.812	1.23
Scaffolding	.810	1.23
Tutor Tell	.614	1.63
Unclear Feedback	.484	2.07
Erroneous Feedback	.791	1.26
Maze		
Fidelity	.818	1.22
Scaffolding	.794	1.26
Tutor Tell	.636	1.57
Unclear Feedback	.493	2.03
Erroneous Feedback	.791	1.26
ASKIT		
Fidelity	.822	1.22
Scaffolding	.823	1.21
Tutor Tell	.648	1.54
Unclear Feedback	.493	2.03
Erroneous Feedback	.791	1.26

Main analysis. To answer research question 4, I conducted hierarchal multiple linear regression. Separate regressions for each of the three outcome variables was done because it was previously noted in the correlation matrix (see Table 12) that all the outcome variables were independent from each other. Thus, it was not necessary to factor in the potential effects of correlated multiple dependent variables. In each of the regression analyses, a hierarchal procedure was employed whereby in the first step the

control variables (i.e., pretest score, race, parents' educational level, and fidelity of implementation) were entered. In the second step, the frequency of scaffolding was entered, since it was the spontaneous instructional behavior with the highest frequency. After that, variables for each of the spontaneous instructional behaviors were entered in the descending order of frequency: tutor tell, unclear feedback, and erroneous feedback. Summaries of the results are presented in Tables 14-16.

Predicting GMRT comprehension scores. The analysis indicated that the full model accounted for 7.1% of the variance in GMRT comprehension scores over and above the impact of race, parents' educational level, pretest score, and fidelity of implementation. Together, race, parents' educational level, GMRT pretest score, and fidelity of implementation accounted for 11.4% of the variance in GMRT posttest scores. However, it can be noted that the overall effect for the four spontaneous tutors behaviors was not statistically significant ($F=1.19$, $p=.332$). An examination of the individual contribution of the predictor variables indicated that none was statistically significant. By examining the change in R^2 , we can understand the contribution of each of the predictors for the variance in GMRT posttest scores. When scaffolding was entered into the model, it did not explain any of the variance in GMRT posttest scores ($\Delta R^2 = .000$) over and above the baseline model. Next, frequency of tutor telling was entered into the model,

and explained 5.7% of the variance in GMRT posttest scores. Unclear feedback explained 1.3% of the variance in student outcome scores on the GMRT, over and above that explained by the baseline model and telling, while 0.2% of the variance can be attributed to the level of erroneous feedback, over and above that explained by the control variables, telling, and unclear feedback. Table 14 details each step of the analysis.

Predicting Maze scores. The overall model with the combined effect of the predictor variables was statistically significant ($F=2.37$, $p=.031$). Overall, the tutors' spontaneous instructional behaviors accounted for 10.7% of the variance in Maze scores over and above the effect of race and parents' educational level, controlling for pretest scores and percentage of fidelity, which together accounted for 18.1% of the variance. Examination of the individual contribution of the predictors indicated that only one variable –scaffolding ($\beta=.272$, $p=.044$) had a statistically significant positive relationship with the comprehension outcome measure. The amount of scaffolding accounted for 7.2% of the variance in Maze posttest scores, controlling for race and parents' educational level, pretest scores and percentage of fidelity. To examine the individual contributions of the other predictor variables, I examined the change in R^2 as each predictor was entered into the model. Telling only accounted for 0.5% of the explained variance above that explained by scaffolding and control variables in the baseline model;

while unclear feedback and erroneous feedback explained 0.8% and 2.2%, respectively.

Thus, in summary, the largest portion of variance in Maze posttest scores explained by tutor behaviors could be attributed to the amount of tutor scaffolding, while the other predictors provided negligible contributions to student outcome scores. Results are summarized in Table 15.

Predicting ASKIT scores. Results showed that the predictors did not significantly predict ASKIT scores ($F=.516, p=.818$). The combined impact of the predictors only accounted for 2% of the variance in ASKIT scores, above that accounted for by the control variables (5%). None of the individual coefficients for the predictors were statistically significant. Of the four predictors, the largest portion of the variance (1.2%) could be attributed to the amount of tutor telling. Both scaffolding (0.6%) and unclear feedback (0.2%) explained a very small portion of the variance in outcome scores on ASKIT, while the amount of erroneous feedback did not explain any of the variance in ASKIT scores ($\Delta R^2 = .000$). Results are summarized in Table 16.

Summary of Findings

Four types of spontaneous tutor behaviors were identified through qualitative coding: (a) scaffolding, (b) telling students the answer, (c) providing unclear feedback, and (d) giving erroneous feedback. Scaffolding was the behavior observed to occur with

the greatest frequency, followed by telling the answer, and providing unclear feedback.

Erroneous feedback occurred least often. Tutors differed in terms of the frequency in which they exhibited each of these behaviors.

Overall, the spontaneous instructional behaviors of tutors accounted for very little variance in student comprehension outcomes. The amount of scaffolding positively influenced Maze scores. Contrary to expectations, the predictors did not substantially account for GMRT scores or ASKIT scores. A detailed look at the contributions of the individual predictors for GMRT Reading Comprehension and ASKIT did not yield significant effects for any particular predictor on any of the outcome variables.

Table 14
 Hierarchical Regression Analysis for Predicting GMRT Comprehension Scores

	β	t	p	F	sig.	R ²	ΔR^2	ΔF	sig ΔF
Step 1				1.635	.180	.114	.114	1.635	.180
Race	0.72	.535	.595						
Parents Educ'l Level	0.56	.422	.675						
GMRT Pretest Score	.275	2.067	.044						
Fidelity	.151	1.144	.258						
Step 2				1.285	.285	.114	.000	.013	.910
Race	.072	.533	.596						
Parents Educ'l Level	.058	.428	.670						
GMRT Pretest Score	.276	2.050	.046						
Fidelity	.153	1.138	.260						
Scaffolds	0.15	.144	.910						
Step 3				1.683	.145	.171	.057	3.371	.072
Race	.016	.114	.910						
Parents Educ'l Level	-.003	-.024	.981						
GMRT Pretest Score	.232	1.735	.089						
Fidelity	.194	1.458	.151						
Scaffolds	-.010	-.072	.943						
Tutor Tell	-.259	-1.836	0.72						

Table 14 continued
 Hierarchical Regression Analysis for Predicting GMRT Comprehension Scores

	β	t	p	F	sig.	R ²	ΔR^2	ΔF	sig ΔF
Step 4				1.543	.176	.184	.013	.754	.390
Race	.004	.026	.979						
Parents Educ'l Level	-.025	-.180	.858						
GMRT Pretest Score	.249	1.841	.072						
Fidelity	.145	1.005	.320						
Scaffolds	.037	.257	.798						
Tutor Tell	-.184	-1.105	.275						
Unclear Feedback	-.153	-.868	.390						
Step 5				1.338	.249	.185	.002	.102	.751
Race	.004	.026	.979						
Parents Educ'l Level	-.023	-.160	.874						
GMRT Pretest Score	.249	1.824	.074						
Fidelity	.142	.974	.335						
Scaffolds	.031	.209	.836						
Tutor Tell	-.181	-1.077	.287						
Unclear Feedback	-.132	-.697	.489						
Erroneous Feedback	-.047	-.319	.751						

Table 15
Hierarchical Regression Analysis for Predicting Maze Scores

	β	t	p	F	sig.	R ²	ΔR^2	ΔF	sig ΔF
Step 1				2.819	.034	.181	.181	2.819	.034
Race	-.016	-.124	.902						
Parents Educ'l Level	-.056	-.434	.666						
Maze Pretest Score	.401	3.120	.003						
Fidelity	-.165	-1.298	.200						
Step 2				3.389	.010	.253	.072	4.822	.033
Race	-.012	-.098	.922						
Parents Educ'l Level	-.026	-.205	.839						
Maze Pretest Score	.451	3.577	.001						
Fidelity	-.138	-1.120	.268						
Scaffolds	.276	2.196	.033						
Step 3				2.839	.019	.258	.005	.320	.574
Race	.004	.033	.974						
Parents Educ'l Level	-.008	-.061	.952						
Maze Pretest Score	.461	3.597	.001						
Fidelity	-.150	-1.192	.239						
Scaffolds	.285	2.232	.030						
Tutor Tell	.075	.566	.574						

Table 15 continued
 Hierarchical Regression Analysis for Predicting Maze Scores

	β	t	p	F	sig.	R ²	ΔR^2	ΔF	sig ΔF
Step 4				2.480	.029	.266	.008	.499	.483
Race	.013	.096	.924						
Parents Educ'l Level	.009	.070	.944						
Maze Pretest Score	.458	3.557	.001						
Fidelity	-.113	-.831	.410						
Scaffolds	.250	1.817	.070						
Tutor Tell	.020	.126	.900						
Unclear Feedback	.117	.706	.483						
Step 5				2.372	.031	.288	.022	1.453	.234
Race	.012	.096	.924						
Parents Educ'l Level	.000	.000	1.00						
Maze Pretest Score	.458	3.566	.001						
Fidelity	-.103	-.757	.453						
Scaffolds	.272	1.972	.044						
Tutor Tell	.010	.062	.951						
Unclear Feedback	.043	.245	.807						
Erroneous Feedback	.167	1.206	.234						

Table 16
 Hierarchical Regression Analysis for Predicting ASKIT Scores

	β	t	p	F	sig.	R ²	ΔR^2	ΔF	sig ΔF
Step 1				.904	.446	.050	.050	.904	.446
Race	-.017	-.123	.902						
Parents Educ'l Level	.134	.975	.334						
Fidelity	-.180	-1.22	.189						
Step 2				.756	.559	.056	.006	.347	.558
Race	-.020	-.147	.884						
Parents Educ'l Level	.124	.893	.376						
Fidelity	-.189	-1.379	.174						
Scaffolds	-.081	-.589	.558						
Step 3				.731	.604	.068	.012	.649	.424
Race	-.049	-.340	.735						
Parents Educ'l Level	.096	.669	.507						
Fidelity	-.171	-1.227	.226						
Scaffolds	-.091	-.657	.514						
Tutor Tell	-.117	-.806	.424						

Table 16 continued
 Hierarchical Regression Analysis for Predicting ASKIT Scores

	β	t	p	F	sig.	R ²	ΔR^2	ΔF	sig ΔF
Step 4				.611	.720	.070	.002	.079	.780
Race	-.052	-.360	.720						
Parents Educ'l Level	.088	.599	.552						
Fidelity	-.187	-1.232	.224						
Scaffolds	-.076	-.505	.616						
Tutor Tell	-.093	-.544	.589						
Unclear Feedback	-.052	-.281	.780						
Step 5				.516	.818	.070	.000	.020	.887
Race	-.052	-.357	.723						
Parents Educ'l Level	.087	.584	.562						
Fidelity	-.185	-1.208	.233						
Scaffolds	-.073	-.476	.637						
Tutor Tell	-.094	-.545	.588						
Unclear Feedback	-.062	-.311	.757						
Erroneous Feedback	.022	.143	.887						

Chapter V: Discussion

The purpose of this study was to (1) identify the types of spontaneous instructional behaviors tutors exhibit in response to student errors while implementing a scripted Tier 2 reading intervention, and (2) explore the relationship between those behaviors and student reading comprehension outcomes. Previous literature has substantiated the relationship between instructional behaviors and student achievement (Duffy et al., 1987; Hoffman et al., 1984; Lara & Medley, 1987; Mariage, 1995; Martin et al., 1980; Samph, 1974; Taylor et al., 2000; 2003; 2005); however no research to date has examined the relationship between instructional behaviors and student achievement in the context of a small group reading intervention using a standard protocol approach. Results from the larger intervention study (Ritchey et al., 2012) indicated that students who received the intervention did not outperform students in the control group on most measures of reading growth. The present study was set within the larger intervention study conducted by Ritchey et al. (2012), and focused on lessons in which the tutors taught the students how to identify the main idea of a paragraph. This study provided an in depth analysis of the tutor-student interactions that occurred during the intervention instruction. Examining the types of spontaneous instructional behaviors used by tutors in reaction to student oral responses and examining fidelity of implementation may point to

how to strengthen the intervention applied in the larger study to yield a greater effect.

Results from this study can help guide the design of future standard protocol

interventions, and inform the training of tutors implementing such interventions.

In this chapter, I present a discussion of the findings, situating the results in terms of relevant literature. Following the discussion of the results, I detail the implications of this study, limitations, contribution to the field, and directions for future research.

Types of Spontaneous Instructional Behaviors

Tutors exhibited four types of behaviors: (a) scaffolding, (b) tutor telling, (c) unclear feedback, and (d) erroneous feedback. Each type of behavior is discussed in light of existing literature, including how each behavior was related to student outcomes in the current study.

Scaffolding. Scaffolding was the most prevalent type of spontaneous instructional behavior exhibited by tutors. Scaffolding typically occurred while the tutor was trying to guide the student to identify the main idea of the passage or after a student answered a question incorrectly. Existing research suggests that scaffolding may be of particular importance for struggling learners, since scaffolding effectively reduces the cognitive load, making success more likely, which consequently impacts motivation and increases learning (Hattie & Gan, 2011; Kluger & DeNisi, 1996). Tutors who used scaffolding

broke down complex processes, such as identifying the main idea, into more simplistic steps, which were more easily completed by students. Scaffolding is a type of sustaining feedback, that is, feedback that extends the interaction between teacher and student.

Teachers who use higher rates of sustaining feedback have been associated with superior literacy outcomes for students in previous literature (Hoffman et al., 1984; Martin et al., 1980).

Tutors in this study used nine types of scaffolding: (a) starters, (b) directives, (c) leading, (d) prompts, (e) multiple-choices, (f) reminders, (g) connections; (h) rephrasing, and (i) metacognitive scaffolds. Researchers have identified similar types of scaffolding in previous studies (Many, 2002; Rodgers 1999; 2004). Similarly, Many (2002) identified nine types of scaffolding in a qualitative study of instructional scaffolding (e.g., modeling, supplying, clarifying, assisting, questioning, prompting, focusing attention, encouraging self-monitoring, and labeling-affirming). Several of these behaviors overlap with the types of scaffolding I identified. For example, “encouraging self-monitoring” is similar to metacognitive scaffolding; and prompting, as described by Many, included the behaviors I identified as leading and prompting.

One type of scaffolding behavior both Rogers (2004) and Many (2002) identified was modeling or demonstrating; however, the tutors in this study did not demonstrate any

additional modeling or demonstrating beyond the modeling scripted in the lesson. In fact, the majority of the tutors did not even provide the explicit modeling of the strategy that was included in the script. Previous research has emphasized that explicit modeling is an essential component of effective reading comprehension instruction, particularly for struggling readers (Book et al., 1985; Duffy et al., 1986; Gersten et al., 2001). I had expected that tutors might have used modeling to demonstrate the targeted strategy; however, I could find no instances of tutor modeling in response to students who were having difficulty performing the strategy independently. Tutor training for future interventions should provide explicit emphasis on the importance of modeling when teaching struggling readers. More research is needed to determine why tutors might be reluctant to model.

In this study, the frequency of tutor scaffolding was found to have a significant positive relationship with student outcomes on one measure of reading comprehension (Maze). Scaffolding explained 7.2% of the variance in Maze posttest scores. However, scaffolding did not account for a significant portion of the variance in scores on the two other comprehension outcome measures. Frequency of scaffolding explained no variance for the GMRT Comprehension, which required students to answer multiple-choice questions based on short passages of text. Scaffolding only accounted for an insignificant

0.6% of the variance in ASKIT scores, a comprehension measure which asked students to apply the comprehension strategies they learned to expository texts. It is possible that the type of comprehension assessment was a factor in explaining differences in the findings.

Maze is a curriculum-based measure (CBM) and CBM has been shown to be more sensitive to reading growth than traditional, norm-referenced measures (Deno et al., 2001), such as the GMRT.

Research supports the finding that scaffolding may play a role in student reading achievement. Wharton-McDonald et al. (1988) found that scaffolding was one of five characteristics that distinguished outstanding teachers of literacy from typical teachers. However, Wharton-McDonald et al. did not use student achievement to distinguish the teacher groups, and instead selected teachers based on supervisor nominations. Wharton-McDonald et al. also suggested that future quasi-experimental studies examine the role of each characteristic and its relationship to student achievement, as their study qualitatively described the differences between teachers who were considered to be highly effective and more typical.

According to the qualitative analysis in the current study, those tutors with previous teaching experience exhibited a greater variety in terms of the types of scaffolding they provided. It may be that tutors develop a repertoire of scaffolding

behaviors over time as a result of their experiences interacting with children. This study is the first to consider that variety in scaffolding types and rate of telling may be related to the instructor's level of experience. Follow up studies should systematically examine how novice instructors compare to those with more experience in terms of the type and variety of scaffolds they provide during instruction.

Tutor telling. Another way tutors responded to student errors was by supplying the answer to the student. Telling occurred following one or more attempts at scaffolding, or immediately following the student offering an incorrect answer to a tutor question. According to the qualitative analysis, the tutors who used highest rate of telling behaviors also used fewer scaffolding behaviors, and vice versa. This was not a consistent finding across all tutors, just those at the extremes of telling and scaffolding. It was also noted that telling seemed to be used more often by tutors who had little previous classroom experience and demonstrated fewer types of scaffolds in their interactions. It may be that tutors who do not have a wide variety of scaffolding prompts to draw on have little option beyond simply telling the student the answer. The finding that some instructors use higher rates of terminal feedback compared to sustaining feedback has been established in previous studies (Chinn, Waggoner, Anderson, Schommer, & Wilkinson, 1993; Taylor et al., 2000; 2002). Previous studies have found that higher frequencies of terminal

feedback are related to poor student outcomes. For example, in a study of the instructional interactions of 92 primary grade teachers, Taylor et al. (2000) found that the least effective teachers told their students the correct answer at a significantly higher rate than their more successful peers. Finn and Metcalfe (2010) examined the differential impact of scaffolding versus telling, finding that although scaffolding and telling did not differentially impact performance levels on assessments given immediately following instruction, scaffolding was superior to telling for longer term recall if testing was delayed. Other researchers have also found support favoring scaffolding over simply providing the student with the information. Pinnell, Lyons, DeFord, Bryk, and Seltzer (1994) found that the most effective tutors tended to provide students with prompts, rather than directly providing an answer.

Despite these findings by other researchers, in the current study, the number of times a tutor told a student the correct answer did not have a significant relationship with any of the comprehension outcome measures. Although the effect was not significant, the frequency of tutor telling explained 5.7% of the variance in GMRT posttest scores, in that higher rates of tutor telling had a negative relationship with GMRT scores ($\beta = -.181$; $p = .29$). Telling accounted for a negligible proportion of the variance for the other two

outcome measures (explaining 0.5% of the variance in Maze scores and 1.2% of the variance in ASKIT scores).

It is difficult to draw any substantive conclusions about tutor telling using the findings of this study alone, given the lack of significance. Although previous literature suggests that telling may not be the best choice for tutors in terms of supporting student learning, further investigation is needed to determine the impact of telling on student reading comprehension achievement.

Another interesting aspect of telling is that although some tutors tended to have consistent rates of telling across all of the groups they taught, others seemed to adjust the frequency of telling they did depending on which group of students they were working with. Cole (2006) found that teachers differentiated their interactions with students during oral reading depending on if the student was a novice or fluent reader. Similarly, in a study of teachers' responses to oral reading errors, Mertzman (2008) found that teachers tended to individualize the scaffolds depending on characteristics of the students.

Mertzman found that teachers differed in their rates of interrupting oral reading and the types of scaffolds they used depending on the ethnic or perceived socioeconomic background of the student. Without interviewing teachers, I cannot determine why tutors elected to tell the answer more frequently for some groups of students than others, but

future research should examine how tutors make decisions about the way they react to student errors.

Unclear feedback. The third type of spontaneous instructional behavior tutors exhibited in this study was unclear feedback. Tutors sometimes delivered feedback that was vague or confusing in reaction to student oral responses. This unclear feedback did not definitively communicate if the answer was correct or incorrect, nor did it identify for the student what about the response was erroneous. Teachers who used unclear feedback responded to student errors hesitantly, using ambiguous phrases such as “Kind of,” or “You could say that.” This type of feedback was less common than scaffolding or telling. It is not clear why tutors seemed hesitant to communicate to the students that the given answer was incorrect; however previous literature has suggested this is not uncommon.

In a study of patterns of tutoring discourse, Graesser, Person, and Magliano (1995) described the difficulty tutors had in providing clear feedback in response to student errors as a “tradeoff between the cognitive goal of imparting correct knowledge and the affective goal of building student confidence” (p. 514). Research on politeness and tutoring has demonstrated that tutors tend to deal with situations requiring criticism, disapproval, or disagreement by being intentionally indirect in their communication and using hedging (such as “kind of”) as a coping mechanism (Bell, Arnold, & Haddock,

2009). Students were selected for participation in the intervention based on the probability of risk for reading failure, and likely had experienced some difficulty in reading in the past. Perhaps tutors did not want to further discourage these students by pointing out their errors. Future research should include tutor interviews to determine why tutors may have felt uncomfortable offering more direct feedback, and instead provided unclear feedback. In addition, studies should examine the impact of adding tutor training for the delivery of feedback on student outcomes.

The fact that tutors provided unclear feedback is troubling considering the research on the power of feedback in relation to student outcomes (see Hattie & Timperly, 2007). Hattie and Gan (2011) explain that the power of feedback can be best realized when the feedback is explicit, making the criteria for success transparent for the learner. Pashler et al. (2005) found that student learning was best when the feedback they received clearly communicated the correct answer, in addition to relaying whether their answer was correct or incorrect. Further, previous research has indicated that unclear feedback can be detrimental to students' performance and self-esteem (Hattie & Timperly, 2007). While tutors may have provided this type of feedback to preserve the students' self-esteem, previous research on unclear feedback suggests it may in fact have

the opposite effect. However, without tutor interviews, I cannot determine the reasoning behind the tutors' use of unclear feedback.

The frequency of unclear feedback did not have a significant relationship with any of the outcome variables. Unclear feedback explained 1.3% of the variance in GMRT scores, 0.8% of the variance in Maze scores, and a meager 0.2% of the variance in ASKIT scores. This study is the first to examine the impact of vague or unclear feedback on student reading comprehension achievement and more research is needed to determine how frequently teachers are delivering this type of feedback. If students, particularly those who are struggling, are frequently exposed to unclear feedback during instruction and, as previous studies have suggested, the effect of unclear feedback is less than ideal, it is important to consider the prevalence of this type of feedback in student-teacher interactions. If unclear feedback is determined to be in common use by some teachers or tutors, professional development and teacher training could certainly address more beneficial means of delivering feedback.

Erroneous feedback. The final way tutors reacted to student oral responses during instruction was by giving erroneous feedback. Tutors provided erroneous feedback either by praising an incorrect answer, ignoring an incorrect answer, or correcting a student's answer even though the student's initial response was correct.

Erroneous feedback was infrequently observed across all tutors (0-8 occurrences per lesson), although, one tutor gave erroneous feedback at more than 3 times the frequency of the other tutors in the study (an average of 5 times per lesson).

Despite being highly trained and well-educated, some tutors delivered erroneous feedback to students. Each of the 12 tutors in this study held a bachelors degree, nine were pursuing a master's degree and three were working toward a doctoral degree. The tutors ranged from 0-7 years of prior teaching experience. Previous literature has found that noncertified teachers aides or volunteer tutors can produce effective outcomes for struggling readers (e.g., Juel, 1996; Savage, Carless, & Erten, 2009; Spear-Swerling, 2009). However, Morris (2006) noted that close supervision by more knowledgeable individuals is key in supporting these individuals. Although the tutors in this study were observed by project personnel at least three times during the intervention, more frequent observations of intervention sessions or weekly reviews of the audio recorded lessons would have assisted project investigators in identifying tutors in need of content knowledge, training, or support. The two tutors who delivered the highest rates of erroneous feedback both lacked professional teaching experience and preservice teacher training. In hindsight, these individuals would have benefitted from additional observations or coaching support throughout implementation of the intervention.

However, it should be noted that the tutors in this study received more supervision than would likely be typical in a school environment outside of controlled experimental research.

Without examining the transcripts with the tutors or gathering additional information through interviews, it cannot be determined why tutors would deliver erroneous feedback, although some possibilities can be considered. Perhaps the tutors themselves did not know how to determine the main idea, and consequently were not equipped to deliver accurate feedback to students on performing a task they could not complete themselves. Another possibility to consider is that the tutor was aware that the student was making an error, but thought the student would improve with more practice, as one tutor's response suggested ("We can work on it a little bit more, but that's pretty good."). Another consideration is that tutor training emphasized the standardized delivery of the intervention, so tutors may have felt powerless to make instructional decisions, such as providing additional practice or modeling, for fear it would jeopardize the efficacy of the intervention.

One common type of erroneous feedback tutors used was providing praise for incorrect main ideas that were stated in fewer than 10 words. In considering this pattern, it is possible the tutor wanted to highlight the students' successive approximations toward

performing the strategy. Tutors may have given this undeserved praise because they wanted to avoid the conflict associated with correcting students or felt that criticism might damage the self-esteem of the struggling readers in front of them.

The frequency of erroneous feedback delivered by tutors did not have a significant relationship with any of the student reading comprehension outcomes. Erroneous feedback did not contribute to the explanation of any variance to ASKIT scores, and had a negligible contribution to the explanation of variance for the other measures, explaining only 0.2% of the variance in GMRT scores and 2.2% of the variance in Maze scores.

Previous research has suggested that erroneous feedback can have a detrimental impact on students. For example, Thompson (1999) found that undeserved praise or unwarranted criticism following task performance caused uncertainty and lowered self-esteem in students. Undeserved praise has been identified as leading to students' use of self-handicapping strategies, such as making excuses for their poor performance (i.e., test anxiety) or avoiding practice (Smith, Snyder, & Handelsman, 1982; Thompson & Richardson, 2001). Although erroneous feedback was not related to student outcomes in this study, based on previous literature we can conclude that erroneous feedback would

not be particularly helpful both in regard to improving reading comprehension and fostering self-efficacy.

Fidelity of Implementation

Three aspects of fidelity were identified as frequently omitted by tutors: (a) modeling, (b) discussing the purpose of the strategy, and (c) providing opportunities for practice. Examining patterns of behavior related to the tutors' adherence to the tutoring protocol produced some interesting findings that can inform the designers of future interventions and tutor-trainers, as well as provide considerations for further research.

Modeling. Despite teacher modeling being widely accepted as an effective instructional practice, particularly for struggling readers (Book et al., 1985; Duffy et al., 1986; Duke & Pearson, 2002; Gersten et al., 2001), the majority of the tutors omitted the modeling portion of the scripted lesson. Instead of explicitly modeling the strategy, 83% of the tutors opted to use a “guided practice” approach when discussing the examples intended for tutor modeling in the script. Considering literature supporting cognitive modeling in teaching reading comprehension (Dole, Nokes, & Drits, 2009; Duke & Pearson, 2002), the lack of tutor modeling observed in this study, both in adhering to the scripted lessons and in response to student errors, may explain the lack of significant findings for the intervention itself.

Previous research has indicated that teachers find modeling difficult (Fisher, 2002). Perhaps the tutors felt that students would lose interest if they engaged in too much “teacher talk” by providing metacognitive modeling of the thinking process used while they determined the main idea, and thus opted to involve students in these examples, rather than modeling them. Further research is needed to determine the impact of frequent versus infrequent teacher modeling during reading instruction, as well as investigation into why teachers or tutors may be reluctant to engage in metacognitive modeling.

Providing purpose for strategy. Another aspect of the scripted lessons that was often omitted by the tutors was explaining the purpose of the reading strategy being taught. Although providing the purpose for the strategy (i.e., when and how it is used) is a key element of explicit strategy instruction (Archer & Hughes, 2011; Duke & Pearson, 2002), tutors did not provide the purpose for learning the strategy in 57% of the lessons. Explaining a strategy’s purpose can help students know when they should employ the strategy and provide motivation for learning it (Dole et al., 2009). Results from this study provide direction for improving tutor training. When training tutors to explicitly teach reading comprehension strategies, the rationale for each section of the strategy lesson should be fully explained. Perhaps if tutors understood why discussing the usefulness of

the strategy with the students was so important, fewer tutors would have omitted this portion of the lesson plan.

Opportunities for practice. Another area in which several of the tutors did not meet the fidelity criteria was in providing each student in the group with an opportunity to independently practice the strategy. Although the scripted lesson indicated that each student should practice the paragraph shrinking strategy, tutors provided practice opportunities for every student in the group in only roughly half of the lessons (52%). Research on instruction is clear: ample opportunities to practice and receive corrective feedback are associated with improved academic outcomes (Hattie & Timperley, 2007; Shute, 2008). Again, this is an area in which closer monitoring of the tutors would have been beneficial. Another potential remedy would be the inclusion of a self-monitoring sheet that the tutors could check off during instruction to ensure that each student had an opportunity to practice the strategy themselves.

Summary. Although each of these aspects of fidelity were not individually examined in relation to student outcomes, future research should consider the individual contributions of each of these aspects of fidelity (modeling, providing a purpose for the strategy, and opportunities for practice) and their relationship to student reading comprehension outcomes. Additionally, further research should consider the use of

interviews to explore the motivations behind omitting aspects of scripted instruction.

Understanding why tutors or teachers elect to skip certain portions of instruction could inform the development of interventions as well as provide information about areas of need for tutor training or professional development.

Other Considerations

Intervention dosage. The spontaneous instructional behaviors that tutors engaged in did not explain a significant portion of the variance for two of the outcome measures – GMRT Comprehension and ASKIT. Only one behavior –scaffolding– significantly contributed to the variance in Maze posttest scores, controlling for student demographic variables, pretest performance and tutor fidelity of implementation. It is not clear why the frequency of scaffolding did not have a significant relationship with the other two reading comprehension outcome measures. As described in Chapter 3, the larger study that examined the effectiveness of the intervention itself did not find significant differences for students who received intervention compared to those who did not on most measures of reading growth. It could be that the intervention (i.e, 16 hours of intervention over 8-10 weeks) was not delivered with enough intensity to produce substantial gains. Vaughn, Denton, and Fletcher (2010) have identified several intervention variables that relate to intensity, including the frequency of sessions (number of times per week), the length of

each session (number of minutes per session), and the duration of the intervention (number of weeks or months for which the intervention is provided). The intervention, which provided the backdrop for this study, was delivered three times per week for 40 minute sessions over 8-10 weeks. It may be unreasonable to expect to be able to detect the influence of these subtle tutor behaviors after such a short time. There are relatively few studies that directly compare effects based on manipulating the intervention dosage (i.e, frequency, length, and duration), and none in the area of reading comprehension. Findings from a review of reading interventions by Vaughn et al. (2010) indicated that it is possible to achieve word-reading gains in a relatively short period of time (8-12 weeks) if the intervention is delivered in frequent, lengthy sessions (approximately 2 hours/day). It may be that meeting 3 days per week for 40 minute sessions over 8-10 weeks was simply not intensive enough to improve reading outcomes, particularly on norm-based measures of reading comprehension.

Measurement. The types of assessments used to measure reading comprehension growth should be considered when interpreting the results from this study. A significant relationship between the tutors' instructional behaviors and student reading comprehension outcomes was only found for one of the three measures –

the Maze, a Curriculum Based Measure (CBM). This may be because CBMs are more sensitive to small changes in student growth than norm-referenced measures, such as the GMRT Reading Comprehension (Martson, Fuchs, & Deno, 1986; Shinn, Deno, & Espin, 2000). This failure to detect small changes in student progress may explain why no significant relationship was detected between student progress on GMRT and the instructional variables of interest. In terms of the ASKIT, failure to account for pretest performance might explain the lack of significant findings for this measure. Pretest scores accounted for between 25% and 45% of the variance in student posttest scores for the two measures that were administered at pretest. It is possible that if pretest scores were available and accounted for, the influence of spontaneous instructional behaviors could have been better detected on the ASKIT.

Another issue related to measurement that may have impacted findings is the restricted range of comprehension outcome scores. When dependent variables have a restricted range detection of relationships between variables is constrained (Hallahan & Rosenthal, 1996). Since the sample population was selected based on subpar performance on the screening measures, the range of scores was restricted. Although there are several ways to statistically correct for restricted range (Wiberg & Sundström, 2009), none were used to correct for the range restriction in this study.

Time limits. During training, tutors were instructed to complete all components within 40 minutes. Adhering to this time limit was a measure of controlling for student outcomes that could be attributed to receiving more or less instruction. Further, the intervention scheduled had been coordinated with each of the classroom teachers so that the participating students missed no more than 45 minutes of instruction on the days they received the intervention, including time needed to transition from the classroom to the location for the intervention lesson. The need to strictly adhere to 40-minute lessons may have influenced how tutors responded to students within the scripted lessons. It was noted in the transcripts that tutors sometimes referred to needing to finish the lesson on time (for example, “We only got so much time.”). The strict time limits of the intervention lessons may also have contributed to some tutors’ high frequency of telling or ignoring errors. The pressure to complete all of the components in the prescribed time may have influenced how the tutors reacted to students.

Management. Lesson transcripts suggest some tutors had difficulties managing the challenging behaviors of the children in their groups. One tutor repeatedly said, “Come on, come on” throughout the lessons. Another made several comments such as “Billy, cut it out. Not another word.” and “Excuse me, what did I say about using such language? What did I say about it? What did I say about it? Answer me. Is it

acceptable?” Later in that same lesson the tutor threatened to send the group down to the principal. The influence of behavior management also may have played a role in influencing how the tutors reacted or responded to students. For instance, a tutor’s choice to ignore incorrect answers or tell students the answer may have been a strategy for maintaining momentum. Breaks in momentum have been described in the classroom management literature as times when teachers are at risk for losing control and undesired student behaviors are likely to escalate (Partin, 1996). It is possible that given the time that was wasted by the students’ off task behaviors and the mandate of completing the lesson in 40 minutes, the tutors felt that telling the students the answer and ignoring student misunderstandings or questions were the best options to keep the lesson moving.

Training for this intervention did not address aspects of student misbehaviors.

When designing future interventions, researchers should consider the inclusion of standardized plans for managing behavior. In addition, coaching support could be put in place for tutors struggling with managing the behavior of specific groups.

Implications

One of the purposes of this study was to further examine aspects of the intervention that could shed light on the findings of original intervention study. Students in the intervention group outperformed those in the control group on only one measure of

reading achievement after participating in the 16 hour intervention. Considering that the intervention was designed based on previously validated methods of teaching reading fluency and comprehension, the investigators were surprised that more dramatic effects were not found. However, upon reviewing the lessons in detail, it seems tutors omitted essential areas of instruction that were included in the scripted lessons. In particular, the lack of strategy modeling and opportunities for student practice may explain why students did not progress as expected as a result of instruction. Future interventions should include closer monitoring of tutors while the intervention is underway, particularly for those tutors who have little or no teaching experience. Tutor training should emphasize the importance of modeling and provide time for tutors to practice modeling during training. Additionally, procedures should be put in place to ensure that each student in the group has ample opportunity to practice the strategy at hand.

This study also suggests that higher levels of scaffolding are related to student achievement on curriculum-based measures (CBMs). Since CBMs are frequently employed to monitor progress and response to instruction within RTI models, differences in the frequencies of scaffolding used during Tier 2 instruction have the potential to influence student responsiveness. Results from this study suggest that instructor differences have the potential to produce meaningful differences in student

responsiveness. If RTI data are to be used to inform decisions about student disability status, tutor differences must be considered. Professional development and training should focus on the power of scaffolding, and instructional coaching and fidelity checks should prioritize scaffolding.

Limitations

The results of this study should be interpreted in light of the limitations of this research. First, this study is limited by possible selection bias and sample size, as this research was conducted in a very small sample of students enrolled in parochial schools. The students in these schools were perhaps not as needy as those in other types of schools, as suggested by their pretest data. Another limitation is that it was a retrospective analysis of existing data, and relied primarily on transcripts of audio recordings of lessons. Video recordings of the lessons would have provided richer information in terms of student behavior and the environment in which the tutoring occurred. Further, video recordings would have allowed analysis of the tutors' physical or gestural prompts, which were not captured via audio. Due to constraints on resources, only two of the 24 lessons were analyzed. Tutors may have exhibited other spontaneous instructional behaviors throughout the intervention while teaching other strategies. In addition, only one

intervention was in place, so it cannot be determined how tutors behaviors may have differed if they were implementing another type of instruction.

Another limitation is that although transcripts were used to examine interactions, the data sources lacked insight from the tutors themselves. Tutor interviews would have enhanced this study by providing an understanding of the tutors' reasoning behind the way they responded to the students. Further, tutor interviews could have provided valuable information about the intervention itself and the challenges they faced implementing the scripted intervention.

Further limitations include lack of consideration for the influence of group size. The intervention was delivered in groups of two and three students, however I did not include group size as an independent variable. However, a post hoc analysis of bivariate correlations between group size and the dependent and independent variables, revealed that only the number of tutor tells had a significant relationship with group size.

A final limitation concerns the coding of complex interaction patterns. Several of the interactions included multiple behaviors. For example, the tutor may have initially responded to an error using scaffolding prompts, however if the student did not respond correctly to these attempts at scaffolding, the tutor may have eventually told the student the answer. This may have resulted in the interaction being coded as three instances of

scaffolding and one instance of telling. On the other hand, another tutor may have reacted to an incorrect answer by immediately telling the student answer, also resulting in one instance of telling. While both of these interactions involve one instance of telling, they are not qualitatively the same. Failure to capture the complexity of these interactions may have influenced the findings of this study. Future researchers should consider coding multiple behavior interactions differently than simple interactions, and examining the influence of these more elaborate behavioral patterns.

Summary

This study contributes to the current knowledge base on how tutors implement scripted interventions and respond to student errors. Findings suggest that scaffolding may play an important role in influencing student achievement on curriculum-based measures of reading comprehension. Although telling, unclear feedback, and erroneous feedback did not have a significant relationship with student outcomes; findings from this study suggest several aspects for designers of future interventions to consider in planning tutor training. The existing literature on the influence of these common teacher behaviors on student achievement is sparse, and this study has illuminated several directions for future research.

Appendix A

Summary of Literature Included in Review

Quantitative Studies							
Study	Purpose	Design	Participants	Student Measures	Teacher Measures	Method of Analysis	Results
Alpert, J.L. (1975)	Investigate the relationship between teacher behavior and student performance. Will increases in "good" teacher behavior result in increased student performance.	Correlational Pretest Posttest Random assignment by school	17 second grade teachers 138 second grade students in 13 schools in NYC	Gates McGinitie.	Observation prior to intervention Teacher Report	ANCOVA	No post-test differences on vocabulary or comprehension measures between the two conditions.
Hoffman, J.V., O'Neal, S.F., Kastler, L.A., Clements, R.O., Segel, K.W., & Nash, M.F. (1984)	Investigated the nature and relationship of teacher-student verbal interaction with student reading achievement levels.	Correlational Pretest Posttest	22 second grade teachers 309 second grade students (N = 152 highest reading group; N = 157 lowest reading group) 10 schools in a south central city	California Achievement Test	Teachers self-selected 5 lessons to audio record. Tapes coded using FORMAS taxonomy (Hoffman & Baker, 1981) Frequency count of teacher verbal behaviors in response to student miscues.	Descriptive Multiple Regression	Negative relationship between terminal feedback & student achievement. Type of miscue and type of teacher feedback related
Lara, A.V. & Medley, D.M. (1987)	Is there a pattern of teacher behavior that is optimal for all students? Are different teacher behavior patterns optimal for different types of learners (high vs low ability)?	Observation Correlational	15 reading teachers in grades 3-8 21 classes	IOWA test of basic skills	Teachers expected gain scores (EGS) were calculated based on pre-post data for class (for both high-achieving & low-achieving students) Frequency counts of 22 teacher behaviors	ANOVA	No one pattern emerged as best for all learners. Gain scores for high-achievers were significantly correlated to (a) teacher demonstrating proper listening skills (b) non-verbal communication skills, and (c) avoiding praise/rewards.

							Low-achieving students showed greatest gains when the teacher (a) gave clear explicit directions and (b) did not allow others to speak.
Martin, J., Veldman, D.J. & Anderson, L.M. (1980)	Investigate the relationship between teacher behaviors and student reading achievement	Observation Frequency Counts	15 first-grade teachers 288 first-grade students	Pretest: Metropolitan Readiness Test Posttest: Metropolitan achievement test	25 teacher behaviors response opportunities per minute how they dealt with correct and incorrect answers.	Multiple Regression	<p>Negative relationship with student achievement: (a) more response opportunities per minute; (b) student calls out; (c) student answers "don't know"; (d) no response; (e) critical feedback; (f) terminating feedback; (g) giving the answer feedback.; and (h) calling on non-volunteers.</p> <p>Positive relationship to achievement (a) answering questions correctly; (b) sustaining feedback; (c) called on to answer questions fewer times; (b) asked non-reading questions (such as comprehension questions or "thought" questions); (c) more positive interactions; (d) fewer "don't know" answers and failure to respond.</p>

<p>Samph, T. (1974)</p>	<p>To determine the extent to which teacher verbal behavior influences the literacy skills of below-average readers.</p> <p>Does teacher style of interaction impact student attitude regarding teacher's fairness, competence, and attractiveness?</p>	<p>Observation Pretest Posttest</p>	<p>Unspecified number of teachers</p> <p>155 sixth-grade students</p>	<p>Metropolitan Achievement Test – Language</p> <p>Pupil Attitude Inventory.</p>	<p>Flander's system of interaction analysis: Teacher talk: 1. accepts feelings, * 2. Praises & encourages, 3. Accepts or uses ideas of student * 4. Asks questions * 5. Lecturing 6. Giving directions 7. Criticizing or justifying authority.</p> <p>*indirect behaviors</p>	<p>ANCOVA</p> <p>Coded teacher verbalizations, classified teachers as "direct" or "indirect"</p>	<p>Controlling for initial achievement posttest language skills revealed a significant difference between direct & indirect teachers. Below average readers scored higher on posttest when matched with indirect teachers.</p> <p>A significant difference was also found for the Pupil Attitude Inventory. Students with indirect teachers had more positive attitudes.</p>
<p>Taylor, B. M., Pearson, P. D., Clark, K., & Wapole, S. (2000)</p>	<p>Determine and explain school-level and classroom-level variables related to student achievement.</p>	<p>Observational Correlational Descriptive</p>	<p>92 teachers in grades K-3</p> <p>368 students in grades K-3</p> <p>14 schools in VA, MN, CO, and CA.</p>	<p>GRADE 1: Letter name ID, blending, segmenting, word reading, QRI-II, fluency, retell</p> <p>GRADE 2/3: QRI-II, fluency (wpm), retelling,</p>	<p>Teachers identified as: (a) Accomplished (A) (b) Moderately Accomplished (M) (c) Least Accomplished (L)</p> <p>Observed 1 hour/month over 5 months.</p> <p>Teacher logs of instruction</p> <p>Questionnaires</p>	<p>Descriptive Statistics ANOVA</p> <p>Transcripts coded for interaction styles: 1.coaching/scaffolding 2.modeling/demonstrating 3. engaging students in recitation 4. Telling 5. Explaining</p>	<p>Accomplished teachers used (a) coaching more often (48% of the time), than M (21%) or L (2%) teachers; (b) engaged in less whole-class instruction, (c) spent more time in small group instruction, (d) had students who spent more time on task & (e) had better home communication. Most accomplished teachers also spent 40 more minutes in literacy instruction.</p> <p>Least Accomplished Teachers (a) used telling significantly more (75%), than M (38%) or</p>

						<p>procedures 6. Discussion (Each teacher was given a "preferred interaction style" based on what they used most frequently)</p>	<p>A (7%). 1st grade students with M or A teachers read on average 19 more wpm than the students with teachers rated L. In grade 2, class ORF averages with A teachers were 10 words higher wpm than teachers rated L, even though at pretest the L class means were 8 words higher per minute. Differences in wpm by teacher were less apparent in grade 3.</p>
--	--	--	--	--	--	------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Appendix A. (con't) Summary of Literature Included in Review

Qualitative Studies							
Study	Purpose	Design	Participants	Student Measures	Teacher Measures	Method of Analysis	Results
Wharton-McDonald, R., Pressley, M., Hampston, J.M. (1998)	To determine the literacy practices of teachers whose students demonstrated the highest levels of achievement.	Ethnography	9 first grade teachers and their students in 4 suburban school districts.	Writing Samples (spelling, conventions, length) Reading levels, as determined by teacher in combination with field notes about books read.	Reading specialists & principal nomination along with student achievement identified outstanding vs. average literacy teachers. 1-2 hr observations conducted twice/month over 7 months. Interviews Artifacts (lesson plans, books, student work)	Analysis and coding of field notes (two researchers simultaneously observed & compared) Ethnographic analysis of interviews	Teachers of highest achieving students (a) balanced authentic reading and writing with explicit instruction, (b) taught with instructional density (multiple goals per lesson), (c) used extensive scaffolding, (d) provided encouragement of self-regulation, (e) integrated reading writing, (f) had high expectations for all students, (g) masterful management, and (h) awareness of the purpose of instructional practices.

Duffy, G.G., Roehler, L.R., & Rackliffe, G. (1987)	To determine the relationship between teacher explanation and student understanding of lesson content.	Descriptive Study Post hoc analysis of a larger study (Duffy, Roehler, & Wesselman, 1985)	7 fifth-grade teachers and 35 students (5 per class of the "lowest" students identified by their teacher)	Post lesson structured interviews recalling lesson material, scored with researcher-created Awareness of Lesson Content Rating Scale.	Researcher-designed Explanation Rating Scale	Analysis of lesson transcripts. Identified Themes Descriptive statistics	Teachers of students whose students scored lower in awareness, engaged in recitation of facts, such as key vocabulary words, and presented reading comprehension strategies as a set of rigid steps. They also tended to use isolated, contrived examples when teaching and asked the students more questions. Teachers whose students demonstrated greater awareness of lesson content modeled the mental process of using the comprehension strategy. These teachers used authentic examples of text and emphasized use of the strategy outside of school. Teachers who elicited better understanding also elaborated on student answers more frequently.
----------------------------------------------------	--------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------	----------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Appendix A. (con't) Summary of Literature Included in Review

Mixed Methods Studies							
Study	Purpose	Design	Participants	Student Measures	Teacher Measures	Method of Analysis	Results
Mariage, T.V. (1995)	Compare the types of teacher talk of more and less effective teachers.	Observational Descriptive	2, 3, 4, 5 15 teachers rank ordered based on student free written recall growth from pre-to-post, Top 5 were "high gainers"; unclear who were low gainers. Three teachers randomly selected from each group	Growth in total ideas recalled	Lesson Transcripts	Descriptive Statistics No tests of significance Analysis of field notes	Teachers who were identified as High Gainers (HG) used more modeling and think-aloud statements, and turned the conversation back to the students more times than Low Gaining teachers (LG). LG teachers used more evaluative statements in response to students (48 vs. 24) and provided less opportunity to extend thinking
Taylor, B. M., Pearson, P. D., Peterson, D. S., & Rodriguez, M. C. (2003)	Investigated the effects of curricular and teaching variables on student reading achievement.	Observational Correlational Descriptive Used stratified random sampling	88 teachers in grades K-5 792 students in grades K-5 9 schools in rural, urban, and small towns in the SE, SW, and Midwestern U.S.	GATES reading comprehension, wrpm, writing first -letter name, PA, Word dictation	Conducted three 1-hour observations over one school year. CIERA Observation Scheme. (Taylor & Pearson, 2000) Coded for instructional variables including student, engagement, grouping, literacy activities, materials used, teacher interaction style	HLM Analysis of field notes	COMPREHENSION Increases in HL questioning and time on task associated with improved reading comprehension scores (+2.5, +2.0 pts, respectively) <i>Negative Effects</i> Increases in time spent in comprehension skill, as opposed to strategy, instruction, and more passive responding associated with lower comprehension scores. 35% of variance between 1st grade teachers 48% of variance attributed between teachers in grades 2-5.

Taylor, B.M., Pearson, P.D., Peterson, D.S., & Rodriguez, M. (2005)	Determine if the professional development associated with the CIERA Framework lead to improved student reading outcomes. Determine school- and classroom-level variables that account for students growth in reading & writing ach.	Observational Correlational	92 teachers and 733 students in grades 2-5. 13 schools in CT, NC, IA, MN, CA.	Reading Fluency (wcpm) Gates McGinitie Reading Test Writing Prompt (scored on a 4-pt rubric) 83% agreement for two scorers on ¼ of samples.	Conducted three 1-hr observations over one school year using CIERA Classroom Observation Scheme (Taylor & Pearson, 2000) Interviews Logs of teacher study groups (School Fidelity 35%)	HLM Growth Curve Analysis Analysis of field notes and interview data	COMPREHENSION 24% of variance between teachers (10 % between schools) Rote comprehension skill instruction negatively related to comprehension growth.
---------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------	--------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------

Appendix B.
Sample Lesson

School	Group	Lesson	Date	Tutor

Lesson 15: Orangutans

Preparation Checklist (initial)

Read instructional plan and text

Gather relevant materials

Write in student names (tutor rating sheet, partners, assessment probes, etc.)

Materials

Repeated reading text (copies for each student), coaching card

Tropical Forest Mammals

Paragraph shrinking chart

Pencils

Objectives

Fluency: repeated reading with a partner

Comprehension: identify main idea

Science: identify how physical and behavioral characteristics help orangutans to survive

Vocabulary: define appearance, suited, prohibit (review protected)

Agenda

1. Fluency: repeated reading with a partner
2. Comprehension: How is paragraph shrinking like retelling? How is it different?
3. Science: How does an orangutan survive in the rain forests?
4. Vocabulary: What do these words mean? appearance, suited, prohibit

√	Step of Instruction	Instructional Sequence	Fidelity
	START AUDIO RECORDER		
Introduction and Agenda (est. 2 minutes)			
	Introduce day by posting and reviewing the agenda	<i>Let's look at the agenda to see what we are learning today.</i>	0 = agenda questions not read 1 = agenda questions read or student/ students read the questions
	Transition to repeated reading	<i>Let's get started with fluency practice.</i>	0 = not present 1 = transition statement made
			section time ____ 0 = > 2 minutes 1 = < 2 minutes
Repeated Reading (est.7 minutes)			
	Introduce repeated reading passage; hand out materials (text and coaching cards)	<i>You and your partner are going to read a short passage about tropical rain forests.</i>	0 = passage not introduced 1 = passage introduced
	Read passage	<i>Listen to me read it first.</i> Rain forests all over the world are in danger. The trees are being cut down. Some are cut for wood to make paper or firewood. Trees are cut down to clear the land. Then the land is used for farms, ranches, or houses. Often a forest is cleared to create a farm. But the farmer can only grow crops there for a few years. Rain forest soil is not good for farming. Living forests hold the soil in place. When the trees are logged, the land is left bare. Rain washes the dirt away. The dirt clogs rivers and it kills the fish.	0 = passage is not read 1 = passage is read but with more than 1 reading error 2 = passage is read with 0 or 1 errors
	Organize partners	Partners (* first coach)	0 = partners are not organized or organization is awkward and takes more than 1 minute 1 = partners are organized in less than 1 minute
		Read with tutor	
		<i>X will read with X and X will read first.</i> <i>X will read with me. Ready? Start reading</i>	
	Student 1 reads for 2 minutes	<i>Ok, switch.</i>	0 = students read for less than one minute 1 = students read for 1 minute to 2 minutes 2 = students read for 2 minutes
	Students switch roles and other student read		0 = students read for less than one minute 1 = students read for 1 minute to 2 minutes 2 = students read for 2 minutes
	Record one student's repeated reading	Trial 1 Rain forests all over the world are in danger. The trees are being cut down. Some are cut for wood to make paper or firewood. Trees are cut down to clear the land. Then the land is used for farms, ranches, or houses.	Miscue Markings omitted = / substitution = / and write miscue above repeated = ____

		<p>Often a forest is cleared to create a farm. But the farmer can only grow crops there for a few years. Rain forest soil is not good for farming. Living forests hold the soil in place. When the trees are logged, the land is left bare. Rain washes the dirt away. The dirt clogs rivers and it kills the fish. (103 words)</p> <p>Trial 2</p> <p>Rain forests all over the world are in danger. The trees are being cut down. Some are cut for wood to make paper or firewood. Trees are cut down to clear the land. Then the land is used for farms, ranches, or houses. Often a forest is cleared to create a farm. But the farmer can only grow crops there for a few years. Rain forest soil is not good for farming. Living forests hold the soil in place. When the trees are logged, the land is left bare. Rain washes the dirt away. The dirt clogs rivers and it kills the fish. (206 words)</p>	<p>0 = tutor does not read with one student 1 = tutor reads with one student</p>								
	Score reading (after session)	<table border="1"> <tr> <td>Name</td> <td></td> </tr> <tr> <td># words attempted</td> <td></td> </tr> <tr> <td># words incorrect</td> <td></td> </tr> <tr> <td># words total</td> <td></td> </tr> </table>	Name		# words attempted		# words incorrect		# words total		<p>0= not scored, not transferred 1= scored and transferred</p>
Name											
# words attempted											
# words incorrect											
# words total											
	Provide feedback.	<p>1. Compliment 2. You got ___ out of ___ words correct today 3. Next time... (critique)</p>	<p>0 = feedback not provided 1 = feedback provided (all 3 components)</p>								
	Transition	<p><i>Great. Now we are going to review our strategy that helps us understand what we read.</i></p>	<p>0 = not present 1 = transition stated</p>								
			<p>tot time sec time</p> <p>steps completed in: 0 = <6 or >9 minutes 1 = 6-8 minutes</p>								
START TEXT READING BY MINUTE 12											
Text Instruction (est. 8 minutes instruction, 20 minutes reading)											
	Tutor will review skill/strategy introduced during previous session.	<p><i>We are going to read about tropical rain forest mammals today. As we read, we are going to paragraph shrink. What are the steps of this strategy?</i> (review strategy chart) Step 1: Read Step 2: <i>What is the person, animal or thing this is about?</i> Step 3: <i>What is the main idea?</i> Step 4: <i>Shrink it! Let's say it in 10 words or less.</i></p>	<p>0 = not present 1= strategy reviewed 2 = strategy reviewed, and chart used</p>								
	Transition to text	<p>Pass out books, preview text</p>	<p>tot time section time</p>								
	pp. 5-6: read to students	<p><i>Listen really carefully as I read because we're going to paragraph shrink when we're done.</i></p>	<p>0 = not present 1 = text is read to students</p>								

		ask 1 student to paragraph shrink the last paragraph on p. 6 STOP at ...rainforests do. (students read the next paragraph)	0 = not present 1 = student s use paragraph shrinking
pp. 6-9: students orally read, paragraph by paragraph	After reading p. 7-the top of 9 REVIEW mammals : <i>We are going to read about a few mammals that live in the tropical rain forests.</i> Q: <i>What are the characteristics of mammals?</i> [a group of animals that have similar characteristic including live birth, hair, warm-blooded, nurse from mothers] finish reading page 9	0 = text not read or read by tutor 1 = students read the text orally mammal 0 = not present 1 = word is reviewed	
p. 10-11 (map)	Q: <i>Let's look at the map. Remember a map is a text feature that helps the reader to learn more from the text, than just reading the words.</i> <i>Where is the equator? Draw a line with your finger. What is the weather like near the equator? (hot and humid, tropical) Where do you live? How does the weather compare where you live?</i>	0 = not present 1 = questions asked 2 = questions asked and answered (with or without additional prompting)	
Use Table of Contents to locate Orangutan section	<i>Now we are going to read about one animal that lives in the tropical rainforest. . Use the table of contents to find the chapter on orangutans. (pg 25)</i> [Review these text features, if needed]	0 = not present 1 = students use TOC to find chapter	
pp. 25-26 : students read silently	Read first sentence, then introduce appearance introduce appearance : <i>The first sentence says that orangutans are humanlike in appearance. Appearance is a word that describes the way something or someone looks. If orangutans are humanlike in their appearance, how do they look or appear to be human? (similar height and weight, similar face). Describe the appearance of another animal that lives in the tropical rainforest? (e.g., toucans, jaguars)</i> Prompt to read silently to learn how an orangutan's appearance is different than other species of monkeys. After students read: Q: <i>How is an orangutan's appearance different than other monkeys?</i> (shaggy hair, red or orange hair, dark skin under hair, bigger than many other monkeys)	0 = text not read or read by tutor 1 = students read the text silently appearance 0 = not present 1 = word is introduced, but definition is imprecise or meaning is not expanded 2 = word is introduced, definition is precise, meaning is extended 0 = not present 1 = questions asked 2 = questions asked and answered (with or without additional prompting)	
pp. 27-28 : read to students	Read text to students, modeling appropriate reading strategies. Introduce suited : <i>The paragraph says an orangutan's body is well-suited or ideal for its way of life. What does that mean? Ideal or suited means that something is perfect for or just right for something else. How is the orangutan's body is well-suited or ideal for its way of life? (curved fingers to</i>	0 = not present 1 = text is read to students suited	

		<p><i>grasp tree branches, long arms help it to swing from tree to tree).Are these physical or behavioral adaptations [physical] How do these features help an orangutan to survive in the rain forest? (help it to swing from tree to tree) How is your body well-suited or ideal for your way of life?(long legs, flat feet for walking, hands that can grasp things, etc.)</i></p>	<p>0 = not present 1 = word is introduced, but definition is imprecise or meaning is not expanded 2 = word is introduced, definition is precise, meaning is extended</p>
	<p>pp. 28-31 : have students orally read, paragraph by paragraph</p>	<p>pg 28 Q: <i>Why do orangutans build a new nest every night? [it moves to other tree to avoid predators] (briefly discuss) Is this a physical or behavioral adaptation? [behavioral] Why?</i></p> <p>Give every student one opportunity to paragraph shrink from pages 28-31.</p> <p>pg 30 Tough to Tackle: unfortunately, international</p> <p>review protect/protected and introduce prohibit: <i>Laws that have been passed to protect orangutans. These laws prohibit people from bringing orangutans into the US so they will not be taken from their natural habitat.. Prohibit means that someone can't or is not allowed to do something or there is a law against doing something. Children are prohibited from driving a car. What other activities are prohibited? How does that protect you or other people?</i></p> <p>pg 31 Tough to Tackle: prohibit (use word chunks strategy here, point out prefix "pro")</p>	<p>Tackling Tough Words</p> <ol style="list-style-type: none"> 1. Chunk the word into small parts 2. Look for parts you know 3. Put it all together. Move your eyes across the word and say it connecting the chunks 4. Read the sentence again to see if it makes sense! <p>0 = text not read or read by tutor 1 = students read the text orally</p> <p>0 = not present 1 = students use paragraph shrinking</p> <p>protect(ed) 0 = not present 1 = word is reviewed</p> <p>prohibit 0 = not present 1 = word is introduced, but definition is imprecise or meaning is not expanded 2 = word is introduced, definition is precise, meaning is extended</p>
	<p>Assessment Probe</p>	<p>As students are reading, ask them to identify the person, animal or thing and main idea; can have students write it on a post it during silent reading</p>	<p>yes = independent maybe = needed a</p>

		Name		prompt no = not able to identify main idea and supporting details
			yes maybe no	
			yes maybe no	
			yes maybe no	
			yes maybe no	
	EXTRA TIME	Read Tapir's chapter Allow students to read silently or orally with a partner Ask students to paragraph shrink, as needed to provide extra practice to specific student.		0 = not present 1 = agenda questions answered
START SUMMARY AND CLOSING BY MINUTE 37				
Summary and Closing (3 minutes)				
	Summary and review	<i>Let's go back to our agenda questions. We are going to review our agenda in a different way today. We are going to read a question and answer to review what we learned today.</i> Read the agenda questions and have students answer each multiple-choice question.		
	Forward look	<i>Great job. Next time we are going to read about other plants and animals in the tropical rain forest.</i>		
	STOP AUDIO RECORDER			

Appendix C.

Intervention Scope and Sequence

Reading Domain	Skill or concept	Day 1	Day 2	Day 3
Week 1				
		Lesson 1	Lesson 2	Lesson 3
Fluency	Repeated reading	Orientation to instructional setting, rules, and procedures	Teach procedure for repeated reading by self: strategies for reading an unfamiliar word (The Understory & The Herb Layer - <i>Life in the Temperate Forest</i>)	Teach procedure for repeated reading with partner: error types, error correction procedure for partner reading (The Shrub Layer, The Canopy, & The Forest Floor - <i>Life in the Temperate Forest</i>)
Vocabulary targets		temperate, ecosystem, hypothesis	expository, narrative, layers	dependence, interdependence, model
Text(s)		<i>Life in the Temperate Forests</i> (Carson, 2003)	<i>Life in the Temperate Forests</i> (Carson, 2003) <i>A River Ran Wild</i> (Cherry, 1992)	<i>Life in the Temperate Forests</i> (Carson, 2003)
Comprehension focus	Text features of expository text (index, glossary, etc.)	Activating background knowledge: forests and animals that live in the forest; read to/with text	Parts of expository text: similarities and differences between narrative and expository text	Parts of expository text, cont. Skimming for key words when answering questions. Identifying where in text answer was found.
Science Concepts		Characteristics and locations of temperate forests	Layers of temperate forest	Building model of temperate forest; interdependence
Materials		REPEATED READING: N/A BOOKS: <i>Life in the Temperate Forests</i> CHART: Text Features	REPEATED READING: The Understory, Herb Layer BOOKS: <i>Life in the Temperate Forests</i> CHART: Text Features	REPEATED READING: Shrub Layer, Forest Floor, Canopy BOOKS: <i>Life in the Temperate Forests</i> CHART: Text Features

Week 2				
		Lesson 4	Lesson 5	Lesson 6
Fluency	Repeated reading	Read by self (Forests - <i>Temperate Forest Mammals</i> , ch. 1)	Read with a partner (What do Wolves Look Like? - <i>Prowling Wolves</i>)	Read by self (All About Deer - <i>All About Deer</i>)
Vocabulary targets		endangered; camouflage, carnivore	slender, tracks, predator/prey (review camouflage)	herbivore, herd/pack, adapt
Text(s)		<i>Prowling wolves</i> (George, 2005)	<i>All About Deer</i>	<i>All About Deer</i>
Comprehension focus	Previewing and Tackling Tough Words	Previewing: previewing expository texts; using text features such as headings as a previewing strategy Introduce strategy for decoding unfamiliar words within text.	Previewing new texts Decoding and using comprehension monitoring to read unfamiliar words.	Reviewing Steps of previewing. Decoding and using comprehension monitoring to read unfamiliar words.
Science Concepts	Interdependence	Interdependence within ecosystem (endangered species) Wolves: interdependence within packs	Interdependence within an ecosystem. Predator – Prey relationships.	Characteristics of deer that help them survive
Materials		REPEATED READING: Forests <i>BOOK:</i> <i>Prowling Wolves</i> CHART: Previewing & Tackling Tough Words ADDITIONAL: Word Cards	REPEATED READING: What do Wolves Look Like? <i>BOOK:</i> <i>All About Deer</i> CHART: Previewing & Tackling Tough Words ADDITIONAL: Word Cards	REPEATED READING: All About Deer <i>BOOK:</i> <i>All About Deer</i> CHART: Previewing & Tackling Tough Words ADDITIONAL: Word Cards

Week 3				
		Lesson 7	Lesson 8	Lesson 9
Fluency	Repeated reading	Read with a partner (Antlers – <i>All About Deer</i>)	Read by self (Raccoons – <i>Temperate Forest Mammals</i>)	Read with a partner (Beavers – <i>Temperate Forest Mammals</i>)
Vocabulary targets		mammals, markings, physical adaptation	lodge, protect/protection, behavior (review adaptation)	membrane, hibernate, echolocation
Text(s)		<i>Temperate Forest Mammals</i> (Landau, 1996) chapter (<i>Raccoons</i>)	<i>Temperate Forest Animals</i> (Landau, 1996) chapter (<i>Beavers</i>)	<i>Zippering, Zapping, Zooming Bats</i> (Earle, 1995) p. 1-22
Comprehension focus	Comprehension monitoring	Comprehension Monitoring: Stopping after reading a section of text to monitor for sense.	Comprehension Monitoring: Stopping after reading a section of text to monitor for sense. Fix up strategies.	Comprehension Monitoring: Stopping after reading a section of text to monitor for sense. Fix up strategies.
Science Concepts	adaptation and survival	physical adaptations of raccoons	physical and behavioral adaptations of beavers	Special features of bats and how features help bats survival
Materials		REPEATED READING: Antlers BOOKS: <i>All About Deer</i> <i>Temperate Forest Mammals</i> CHART: Monitoring	REPEATED READING: Raccoons BOOK: <i>Temperate Forest Mammals</i> CHART: Monitoring ADDITIONAL: Beaver Diagram	REPEATED READING: Beavers BOOK: <i>Zippering, Zapping, Zooming Bats</i> CHART: Monitoring

Week 4				
		Lesson 10	Lesson 11	Lesson 12
Fluency	Repeated reading	Read by self (Bat Wings - <i>Zipping, Zapping, Zooming Bats</i>)	Read with a partner (What Are Bats? - <i>Zipping, Zapping, Zooming Bats</i>)	Read by self (Owls: Hunting – <i>Owls</i>)
Vocabulary targets		species, habitat, disturb	roost, raptor, hollow	digest, enemy, pollute/pollution
Text(s)		<i>Zipping, Zapping, Zooming Bats</i> (Earle, 1995) p. 23-31	<i>Owls</i> (pp. 1- 13)	<i>Owls</i> (pp. 15-24)
Comprehension focus	Summarizing Main idea	Retelling	Retelling to monitor for comprehension.	Retelling to monitor for comprehension.
Science Concepts	Classification: birds	Bats: special features	Special features of owls and how features help owls survival	Owl Pellet investigation
Materials		REPEATED READING: Bat Wings <i>BOOK: Zipping, Zapping, Zooming Bats</i> CHART: Retelling	REPEATED READING: What are Bats? <i>BOOKS: Owls Zipping, Zapping, Zooming Bats</i> CHART: Retelling	REPEATED READING: Owls: Hunting <i>BOOK: Owls</i> CHART: Retelling ADDITIONAL: Materials for owl pellet dissection (kit) Investigation sheet (copies for each student) Skeleton outlines sheet

Reading Domain	Skill or concept	Day 1	Day 2	Day 3
Week 5 (Animals in the Rain Forests)				
		Lesson 13	Lesson 14	Lesson 15
Fluency	Repeated reading	Read with a partner (Life in a Forest - <i>Temperate Forest Mammals</i>)	Read by self (Layers of the Tropical Rain Forest)	Read with a partner (Endangered Forests)
Vocabulary targets		tropical, humid, diverse (review species, ecosystem)	epiphytes, pollinate, fungi/fungus (review camouflage)	appearance, suited, prohibit (review protected)
Texts		<i>Tropical Rain Forests (Sayre, 2002)</i> (pp. 1-13)	<i>Tropical Rain Forests (Sayre, 2002)</i> (pp. 14-	<i>Tropical Forest Animals (Orangutan chapter)</i>
Comprehension focus	main idea: paragraph shrinking	Review: steps and process for paragraph shrinking	Guided practice: paragraph shrinking	Independent/partner practice: paragraph shrinking
Science Concepts	Rain forests- characteristics and overview of animals that live in rain forest	Characteristics and parts of rain forest (canopy, understory, etc.);	Characteristics of rain forests; bromeliad	Animals in the monkey family (orangutans)
Materials		REPEATED READING: Life in a Forest BOOKS: Tropical Rain Forests (Sayre) CHART: Paragraph Shrinking ADDITIONAL: Inflatable Globe Pencils with Erasers Layers of rain forest worksheet Crayons	REPEATED READING: Layers of the Tropical Rain Forest BOOKS: Tropical Rain Forests (Sayre) CHART: Paragraph Shrinking ADDITIONAL: Spray bottle with water Plastic Frogs Bromeliad Plant Metal Tray	REPEATED READING: Endangered Forests BOOKS: Tropical Rainforest Mammals CHART: Paragraph Shrinking

Week 6				
		Lesson 16	Lesson 17	Lesson 18
Fluency	Repeated reading	Read by self Choice: Orangutans Tropical Forests	Read with a partner Tamarins	Read by self Choice: Howler Monkeys What Tamarins Eat
Vocabulary targets		define territory, scent, classify (review species, endangered)	omnivore, opposable, re-introduce/re-introduction	arboreal, vertebrae, sloth (i.e., laziness, multiple meanings) (review habitat)
Text(s)		<i>Tropical Forest Animals (howler monkey chapter)</i> <i>Tamarins</i>	<i>Tamarins</i>	<i>Sloths</i>
Comprehension focus	Main idea and supporting details: Paragraph Shrinking with supporting details	Review: steps and process for paragraph shrinking; using supporting details	Guided practice: paragraph shrinking with supporting details	Independent/partner practice: paragraph shrinking with supporting details
Science Concepts	Adaptations of Animals in the rain forest	howler monkeys, tamarins; how tamarins have adapted to life in rain forest	how tamarins have adapted to life in rain forest	Characteristics of sloths; how sloths have adapted to life in rain forest
Materials		REPEATED READING: Orangutans Tropical Forests BOOKS: <i>Tropical Forest Animals</i> <i>Tamarins</i> CHART: Paragraph Shrinking	REPEATED READING: Tamarins BOOKS: <i>Tamarins</i> CHART: Paragraph Shrinking ADDITIONAL: Text examples	REPEATED READING: Howler Monkeys What Tamarins Eat BOOKS: <i>Sloths</i> CHART: Paragraph Shrinking

Week 7				
		Lesson 19	Lesson 20	Lesson 21
Fluency	Repeated reading	Read with a partner <i>Sloths</i>	Read by self- Choice <i>Are Sloths Endangered?</i> <i>How Sloths Move</i>	Read with a partner <i>Meet the Jaguar</i>
Vocabulary targets		home range, preserve, fork (review herbivore)	carnivore, muscular, poachers (review carnivore)	reptile, grasslands, scutes/scales
Text(s)		<i>Sloths</i>	<i>Pouncing Jaguars</i>	<i>Anacondas</i>
Comprehension focus	Question Answer Relationships	Explicit teaching/mini lesson: Right There Questions	Explicit teaching/mini lesson: Think and Search Questions	Explicit teaching/mini lesson On My Own Questions
Science Concepts	Interdependence	Adaptation and dependence of sloths	Characteristics of jaguars; role of jaguars in the food chain	Part 1 of anacondas
Materials		REPEATED READING: <i>Sloths</i> BOOKS: <i>Sloths</i> CHART: QAR ADDITIONAL: fork word card, Questions 1, 2, 3	REPEATED READING: <i>Are Sloths Endangered?</i> <i>How Sloths Move</i> BOOKS: <i>Sloths</i> <i>Pouncing Jaguars</i> <i>Tropical Forest Mammals</i> CHART: QAR	REPEATED READING: <i>Meet the Jaguar</i> BOOKS: <i>Anacondas</i> <i>Pouncing Jaguars</i> CHART: QAR TTW Previewing

Week 8				
		Lesson 22	Lesson 23	Lesson 24
Fluency		Read by self –Choice <i>Anacondas</i> <i>Jaguar Enemies</i>	Read with a partner <i>How Anacondas Hunt</i>	Read by self Frogs in Danger
Vocabulary targets		hinge, molt, scent, cold-blooded/warm-blooded (review digest, carnivore)	amphibian, metamorphosis, suction	review words
Text(s)		<i>Anacondas</i>	<i>Watching Tree Frogs in South America</i>	Review activity – food chain and interdependence
Comprehension focus	QAR: Asking Questions & Review	Explicit teaching/mini lesson: Author & Me	Reviewing all strategies	Reviewing all strategies & content
Science Concepts	Classification (reptiles, amphibians)	Part 2 of anacondas <i>Anacondas</i> ; classification of reptiles	Amphibians (tree frogs); classification of amphibians	interdependence: food chain
Materials		REPEATED READING: <i>Anacondas</i> <i>Jaguar Enemies</i> BOOKS: <i>Anacondas</i> CHART: QAR TTW Previewing	REPEATED READING: <i>How Anacondas Hunt</i> BOOKS: <i>Watching Tree Frogs in South America</i> CHART: QAR TTW Paragraph Shrinking Monitoring	REPEATED READING: Frogs in Danger BOOKS: <i>Watching Tree Frogs in South America</i> ADDITIONAL: Game board Animals Dice Questions

Appendix D.

Code Book

Code	Explanation
I-Tell	After an error, the tutor tells the students the correct answer.
I-Terminal	Tutor gives no response to an error or calls on someone else. Tutor indicates that it is wrong “No”
I-Hint	The tutor gives the students a hint (task-focused, not how to use the strategy in the future, may include it starts with X)
I-Direct	Directs student to strategy or do something specific.
I-Direct-Text	Directs student to refer to text (Point to where you found the answer)
I-Direct-reread	Directs student(s) to reread the text.
I-Model	Models how to do it correctly. Differs from tell because modeling would include explicit instructions/think aloud.
I-explains	Tutor explains or clarifies a misunderstanding. May occur after a Tell.
I-repeatQ	Tutor repeats the question using the same words that they used the first time.
I-rephraseQ	Tutor rewords or rephrases the question in a way that remains true to the original question.
I-asifC	Answer is incorrect, but tutor acts as if it is correct. (tutor error)
I-Q	The student made an error, tutor repeats back, questioning the student.

Correct Answer

Code	Explanation
C-Repeat	Tutor repeats correct answer in the same words the student used.
C-Restate	Tutor rewords the correct answer using slightly different words than the student used.
C-Expand	Tutor expands on or explains more about the correct answer.
C-specific_praise	Tutor lends specific praise regarding the strategy the student used to get the correct answer.
C-Terminal	Tutor says, “Good” “Okay” or the like to indicate that the answer was correct, but does not do anything else.
C-Ignore	Tutor asks another question or does not respond to the student’s correct answer. Moves on/ignores (no verbal response)

Questions

Code	Explanation
Q-clarify	Tutor asks a question to clarify a student’s answer.
Q-explain	Tutor asks the student to explain how they got their answer.
Q-script	Tutor asks a question that is stated in the script.
Q-rhetorical	Tutor asks a rhetorical question. (Typically used for behavior management)
Q-rephrase	Use when questions are rephrased before a student attempts to answer.

Code	Explanation
CONNECT	Tutor makes a connection to a previous lesson, earlier within the same lesson, or prior knowledge (not included in the script)
MODEL (not in script)	Tutor infuses additional modeling not suggested in the script.
META	Tutor does metacognitive modeling –modeling the thinking process they use while they model the skill.
SCAF	Tutor scaffolds by asking questions or prompts in an effort to lead student to answer, Scaffold: a support that bridges the gap between a student's current understanding and the intended outcome. May be a series of questions that lead the student to the answer, or prompt(s) that guide the student to complete the task themselves.
CONF	Confusing. Use anytime the tutor gives unclear instructions, explanations.
SCRIPT	Tutor reads from script (may use slightly different wording, but remains true to the intent of the script).
MAN	Tutor addresses Management/Behavior
Read	Tutor Reads
COACH	Provides coaching while student works.
DIRECT	Gives directions
unclear_feedback	Tutor gives unclear feedback (not clear that students are correct/incorrect)
Response	Tutor responds to student, but no code
Tell	Tutor answers their own question
SUMM	Tutor summarizes what has been read.

*Appendix E*ASKIT Assessment and Scoring Examples
SEA TURTLES

Reading Comprehension Assessment

Purpose: To assess the extent to which students can identify and apply the following comprehension strategies: previewing, retelling, fix up strategies for words, identifying main idea, identifying other important details.

Text: Protecting Sea Turtles (Sullivan, 2004)

Administration:

- **Audiotape entire assessment in a separate file in your audiorecorder. (Stop audiorecording from individual testing and start a new file. Repeat ID, child’s initials and examiner name in this audiofile.)**
- **Record student responses to questions verbatim.**
- **Check “Listen to Tape” if you need to go back and listen to the tape to fill in portions of the response not accurately recorded.**
- **After the test session is finished, check items in the scoring checklist that correspond to student’s response. Do not check off items during testing.**
- **File protocol separately in designated folder in black cabinet.**
- **For all items, allow students up to 15 seconds to respond before moving to the next item.**

Explain activity to student.

In this next activity, I will ask you to read an information book, and I will ask you to stop as you read to tell me what you read and to answer questions about the book. I will also ask you to think about what you would tell other fourth graders to do if they want to be good readers. Let’s begin.

Place closed book in front of student and say: *This is the book we are going to read. Please keep it closed for now.*

1a. What would you tell other fourth graders to do before they begin reading if they want to be good readers? When student finishes response, prompt one time by saying: *Can you think of anything else these students should do before reading?*

Answer	Checklist
<p>Prompt: <i>Is there anything else these students should do before reading?</i></p> <p>Need to listen to tape.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> look at pictures or illustrations <input type="checkbox"/> look at the table of contents <input type="checkbox"/> read the title <input type="checkbox"/> read captions <input type="checkbox"/> think about what you know about the topic <input type="checkbox"/> look through book <input type="checkbox"/> other: _____ <input type="checkbox"/> other: _____

1b. You said the students should (insert 1-2 of student's answers). **Why do good readers do these things before they start reading?**

Answer	Checklist
<p>Need to listen to tape.</p>	<ul style="list-style-type: none"> <input type="checkbox"/> to get ready to read <input type="checkbox"/> to learn what you will read about <input type="checkbox"/> to make predictions about what you might learn <input type="checkbox"/> other _____ <input type="checkbox"/> other: _____

2a. Before you start reading, what do you think you might learn in this book? [Hand book to student, students can preview/look through book but should not be verbally prompted to do so.]

Answer	Checklist
<p>___ Check if student previews/looks through book.</p> <p>___ Need to listen to tape.</p>	<p>___ about sea turtles</p> <p>___ where sea turtles live</p> <p>___ what happens to sea turtles</p> <p>___ how people can protect sea turtles</p> <p>___ other</p> <hr/> <p>___</p> <p>___ other</p> <hr/> <p>___</p>

2b. Why do you think you will learn this?

Answer	Checklist
<p>___ Need to listen to tape.</p>	<p>___ used text features</p> <p>___ read table of contents</p> <p>___ looked at pictures</p> <p>___ used background knowledge</p> <p>___ made a guess</p> <p>___ other</p> <hr/> <p>___ other</p>

Okay, open the book to p. 3. Remember that I am going to ask you some questions about what you read. Please read this page out loud to me. Record errors using CBM rules including 3 sec rule if students does not read a word. Do not time the student's reading.

Sea turtles are gentle creatures that live in oceans all over the world. Most sea turtles spend their whole lives in the water. The only time most sea turtles go to shore is to lay their eggs.

of errors _____

Like many wild animals, sea turtles need protection. Why? Let's read more about sea turtles.

At the end of p. 3 ask the student to stop. **Please stop reading here.** Cover text with blank paper.

3. Now tell me what you just read. [If student does not respond, say **That's okay. We will keep reading. Remember that I will ask you about what you have just read.** Record NR in the answer space]

Answer	Checklist
<p>Need to listen to tape.</p>	<p><input type="checkbox"/> live their whole lives in ocean/water</p> <p><input type="checkbox"/> live in oceans</p> <p><input type="checkbox"/> lay eggs on shore</p> <p><input type="checkbox"/> gentle creatures</p> <p><input type="checkbox"/> need protection</p> <p><input type="checkbox"/> other _____</p> <p>_____</p> <p><input type="checkbox"/> other _____</p>

Okay, turn to page 4 and keep reading. Record errors using CBM rules including the 3 sec rule. Do not time.

There are seven different **species**, or types of sea turtles. Most sea turtles have hard shells on their backs. The only sea turtle that doesn't have a hard shell is the leatherback turtle. It has a leathery skin instead of a shell.

of errors _____

At the end of p. 4, ask the student to stop reading. *Please stop reading here.*

4a. Look at this word, species. [Point to species]. *Why is the word species in bold?*

Answer	Checklist
<p>____ Need to listen to tape.</p>	<p> <input type="checkbox"/> hard word <input type="checkbox"/> important word <input type="checkbox"/> may need it explained <input type="checkbox"/> in the glossary <input type="checkbox"/> vocabulary word <input type="checkbox"/> teacher might ask about it <input type="checkbox"/> other _____ <input type="checkbox"/> other _____ </p>

4b. If the 4th grade students you are helping didn't know what species means [Point to species], what would you tell them to do?

Answer	Checklist
<p>Need to listen to tape.</p>	<p> <input type="checkbox"/> ask teacher <input type="checkbox"/> read more sentences <input type="checkbox"/> use glossary/dictionary <input type="checkbox"/> other _____ _____ </p>

5a. This word, leatherback, might be tricky for someone to read. [Point to leatherback.] If the 4th graders you're helping don't know how to read this word, what would you tell them to do?

Answer	Checklist
<p>Need to listen to tape.</p>	<p> <input type="checkbox"/> sound it out <input type="checkbox"/> find parts you know and put them together <input type="checkbox"/> keep reading <input type="checkbox"/> ask the teacher <input type="checkbox"/> tell him/her <input type="checkbox"/> skip it <input type="checkbox"/> other _____ _____ </p>

5b. If student says, “I will tell him the word” or “ask the teacher” for Item 5a, prompt: ***If you wanted them to figure it out on their own, what would you tell them?*** Record NA if this question does not need to be administered.

Answer	Checklist
<p>Need to listen to tape.</p>	<p> <input type="checkbox"/> sound it out <input type="checkbox"/> find parts you know and put them together <input type="checkbox"/> keep reading <input type="checkbox"/> other _____ </p>

Okay, now read page 6 to me out loud. After you read I will ask you to tell me the main idea and details. Remember, the main idea is the most important information. Record errors using CBM rules including the 3 sec rule. Do not time student’s reading.

Most sea turtles return to the same beach where they were born to lay their eggs. The mother sea turtle uses her flippers to move across the sand and dig a hole. This hole will be the nest.

of errors _____

Most sea turtles lay about 100 eggs at a time. When the mother finishes laying the eggs, she covers the hole with sand. Then the mother pushes herself back to the ocean and swims away.

At the end of p. 6 ask the student to stop. ***Please stop reading here.*** Cover text with blank paper.

6a. What is the most important information you just read?

Answer	Checklist
<p>___ Need to listen to tape.</p>	<p>___ return to same beach ___ return to same beach to lay eggs ___ mother makes nest in sand ___ mother makes a hole in the sand ___ a hole is the nest ___ lay 100 eggs at a time ___ mother covers hole with sand ___ mother goes back to ocean ___ other _____ _____ ___ other _____</p>

6b. Did students count number of words in their response? [Using fingers or counting orally]

Circle: Yes No How many words? _____

6c. Now tell me the details, or the information that supports the main idea.

Answer	Checklist
<p>___ Need to listen to tape.</p>	<p>___ return to same beach ___ return to same beach to lay eggs ___ mother makes nest in sand ___ mother makes a hole in the sand ___ a hole is the nest ___ lay 100 eggs at a time ___ mother covers hole with sand ___ mother goes back to ocean ___ other _____</p>

Let's keep reading page 8. After you read I will ask you to tell me the main idea and details. OK, read this page to me.

Record errors using CBM rules including the 3 sec rule. Do not time reading.

Baby sea turtles hatch, or come out of their eggs, after about two months under the sand. The bodies of the baby turtles, or hatchlings, are very soft. They must get to the sea quickly. Sea birds and crabs will catch some of the hatchlings.

The hatchlings swim as soon as they reach the water. But they are in danger. Many will become

of errors _____

food for other animals in the sea. Only a few of the hatchlings will grow to be adult sea turtles.

At the end of p. 8 ask the student to stop. *Please stop reading here.* Cover text with blank paper.

7a. What is the most important information you just read?

Answer	Checklist
<p>___ Need to listen to tape.</p>	<p>___ baby turtles hatch ___ baby turtles hatch in 2 months ___ baby turtles are called hatchlings ___ they must get to sea quickly ___ animals/sea birds/crabs will catch some of them ___ hatchlings/baby turtles are in danger ___ animals will eat them ___ only a few will grow to be adults ___ other</p>

7b. Did students count number of words in their response? [Using fingers or counting orally]

Circle: Yes No How many words? _____

7c. Now tell me the details, or the information that supports the main idea.

Answer	Checklist
	<p>___ baby turtles hatch ___ baby turtles hatch in 2 months ___ baby turtles are called hatchlings ___ they must get to sea quickly</p>

<p>___ Need to listen to tape.</p>	<p>___ animals/sea birds/crabs will catch some of them ___ hatchlings/baby turtles are in danger ___ animals will eat them ___ they will be food for other animals ___ only a few will grow to be adults ___ other _____</p>
------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

1

8a. [Remove cover sheet.] *Here is another word that might be tricky for someone to read.* [Point to hatchlings.] *If the fourth graders you're helping don't know how to read this word, hatchlings, what would you tell them to do?*

Answer	Checklist
<p>___ Need to listen to tape.</p>	<p>___ sound it out ___ find parts you know and put them together ___ keep reading ___ ask the teacher ___ tell him/her ___ skip it ___ other _____ _____ ___ other _____</p>

8b. If student says, "I will tell him the word" or "ask the teacher" for Item 8a, prompt: *If you wanted them to figure it out on their own, what would you tell them?* Record NA if this question does not need to be administered.

Answer	Checklist
	<p>___ sound it out</p>

<p>Need to listen to tape.</p>	<p>__ find parts you know and put them together __ keep reading __ other _____ _____</p>
--------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Note to examiner: Skip page 10-11.

Now read page 12. Record errors using CBM rules including the 3 sec rule. Do not time student's reading.

Protecting Sea Turtles

Many species of sea turtles are endangered. This means that there aren't many of these turtles left. People and animals eat turtle meat and turtle eggs. Sometimes turtles are caught by mistake in fishing nets.

of errors _____

At the end of p. 12 ask the student to stop. ***Please stop reading here.***

11. Here is another word that might be tricky for someone who doesn't know what it means. [Point to endangered.] ***If the 4th grade students you are helping didn't know what endangered means, what would you tell them to do?***

Answer	Checklist
<p>Need to listen to tape.</p>	<p>__ asking teacher __ reading more sentences __ using glossary/dictionary __ other _____ _____</p>

Note to examiner: Skip page 13.

Now read the last page. (p. 14) Record errors using CBM rules including the 3 sec rule. Do not time student's reading.

People can protect the land and water where sea turtles live. In some places, people make sure nests are safe until the baby turtles hatch. Learning about sea turtles helps us know how to protect them now and in the future.

of errors _____

At the end of p. 14 ask the student to stop. ***Please stop reading here.*** Cover text with blank paper.

12. What is the most important information you just read?

Answer	Checklist
Need to listen to tape.	<input type="checkbox"/> people can protect sea turtles <input type="checkbox"/> people can make sure nests are safe until babies hatch <input type="checkbox"/> learning about sea turtles helps us know how to protect them <input type="checkbox"/> other _____

Close book and take it away.

13. Now, think about the whole book. What did you learn about sea turtles in this book?

Answer	Checklist
	<input type="checkbox"/> sea turtles are gentle <input type="checkbox"/> sea turtles spend their whole lives in the water <input type="checkbox"/> sea turtles come to shore to lay eggs <input type="checkbox"/> there are many types of sea turtles <input type="checkbox"/> sea turtles lay their eggs in a nest on the beach they were born

ASKIT Scoring Examples

Question: What is the most important information you just read? (Main Idea)

Correct Answer: Baby sea turtles are in danger when they are born

Score	Description	Examples
0	no response or response incorrect information	<i>Turtle, sea turtles. How sea turtles grow that sea turtles try to swim to shore.</i>
1	response is a single detail from text	<i>Seabirds and other kinds of animals eat them. Some turtles won't be able to grow up to adult turtles because other animals will say umm. That's yummy food. [second part of response is not from text] birds or crabs come and eat the hatchlings</i>
2	response is a general statement about what was read OR response includes more than one detail from text and does not include the main idea (that includes a who and a what	<i>About how their baby sea turtles hatch. When sea turtles are born they usually very few of them get to the water</i>
3	Response includes main idea, may also include other details (main idea has who/what and the most important thing) OR includes a synthesis of the information	<i>That baby turtles are very soft and in danger of other animals coming to eat them. The baby sea turtles are endangered. About how sea turtles are in danger when they hatch That after they are hatch they have to go right to the water but they are in danger because some of the other sea creatures maybe eaten them for lunch and some may stay alive and become adults.</i>

Appendix F

Fidelity Checklist

Lesson 14

	Reviews paragraph shrinking strategy
	Discusses reason for learning strategy (i.e., how strategy helps you as a reader)
	Tutor models strategy
	Tutor provides guided practice of strategy
	Each student has 1 opportunity to practice the strategy
	Vocabulary – Epiphyte. Definition precise, meaning extended
	Vocabulary – Pollinate. Definition precise, meaning extended
	Vocabulary – Fungi. Definition precise, meaning extended
	Reviews camouflage
/9	TOTAL

Lesson 15

	Reviews paragraph shrinking strategy
	Each student has 1 opportunity to practice paragraph shrinking strategy
	Reviews characteristics of Mammals
	Students use table of contents to locate chapter on orangutans
	pp. 25-26 : students read silently
	Vocabulary – Appearance. Definition precise, meaning extended
	Vocabulary – Suited. Definition precise, meaning extended
	Vocabulary – Prohibit. Definition precise, meaning extended
	Tackling tough word strategy is practiced
	Vocabulary – Protect is reviewed
/10	TOTAL

References

- Aarnoutse, C., VanLeeuwe, J., Voeten, M., & Oud, H. (2001). Development of decoding, reading comprehension, vocabulary and spelling during the elementary school years. *Reading and Writing, 14*(1), 61-89.
- Allen, S. A. (1966) The effects of verbal reinforcement on children's performance as a function of the type of task. *Journal of Experimental Child Psychology, 3*, 57-73.
- Allen, S. A., Spear, P. S., & Lucke, J. R. (1971). Effects of social reinforcement on learning and Retention in Children. *Developmental Psychology, 5*(1), 73-80.
- Alpert, J.L. (1975). Teacher behavior and pupil performance: A study of the mediation of Pygmalion effects. *Journal of Educational Research, 69*, 53-57.
- Anfara, V. A. (2006). Utilizing mixed methods in middle grades research. *Middle Grades Research Journal, 1*, 15–31.
- Archer, A., & Hughes, C. A. (2011). *Explicit instruction: Efficient and effective teaching*. New York, NY: Guilford Publications.
- Bakker, J., Denessen, E., & Brus-Laeven, M. (2007) Socio-economic background, parental involvement and teacher perceptions of these in relation to pupil achievement. *Educational Studies, 33*, 175-190.
- Bali, V. A., & Alvarez, R. (2004). The Race Gap in Student Achievement Scores: Longitudinal Evidence from a Racially Diverse School District. *Policy Studies Journal, 32*(3), 393-415. doi:10.1111/j.1541-0072.2004.00072.x
- Bandura, A. (1969). Social Learning Theory of Identification Processes. In D.A Goslin (Ed.), *Handbook of socialization theory and research* (pp 213-262). Skokie, IL: Rand McNally.

- Bandura, A. (1977) Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review*, 84(2), 191-215.
- Bandura, A., & Barab, P. (1971). Conditions governing nonreinforced imitation. *Developmental Psychology*, 5(2), 244-255.
- Barton, P.E., & Jenkins, L. (1995). *Literacy and dependency: The literacy skills of welfare recipients in the United States*. Princeton, New Jersey: Educational Testing Service.
- Bell, D. C., Arnold, H., & Haddock, R. (2009). Linguistic Politeness and Peer Tutoring. *Learning Assistance Review*, 14, 1, 37-54.
- Berkeley, S., Bender, W.N., Peaster, G.L., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities*, 42(1), 85–95.
- Boardman, A. G., Roberts, G., Vaughn, S., Wexler, J., Murray, C. S., & Kosanovich, M. (2008). *Effective instruction for adolescent struggling readers: A practice brief*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Book, C. L., Duffy, G. G., Roehler, L. R., Meloth, M. S., & Vavrus, L. G. (1985). A study of the relationship between teacher explanation and student metacognitive awareness during reading instruction. *Communication Education*, 34(1), 29.
- Borko, H., & Shavelson, R. (1990). Teachers' decision making. In B. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 311-346). Hillsdale, NJ: Erlbaum.
- Bradley, R., & Corwyn, R. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53(1), 371.

- Brantlinger, E., Jimenez, R., Klingner, J., Pugach, M., & Richardson, V. (2005). Qualitative studies in special education. *Exceptional Children, 71*(2), 195-207.
- Brooks-Gunn, J., Klebanov, P. & Duncan, G.J. (1996). Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment and maternal characteristics. *Child Development, 67*, 396-408.
- Brown, K., Morris, D., & Fields, M. (2005). Intervention after grade 1: Serving increased numbers of struggling readers effectively. *Journal of Literacy Research, 37*(1), 61-94.
- Burns, M. K., Appleton, J. J., & Stehouwer, J. D. (2005). Meta-analytic review of responsiveness-to-intervention research: Examining field-based and research-implemented models. *Journal of Psychoeducational Assessment, 23*, 381-394.
- Burns, M. K., Jacob, S., & Wagner, A. (2008). Ethical and legal issues associated with using response-to-intervention to assess learning disabilities. *Journal of School Psychology, 46*, 263-279.
- Caracelli, V.J. & Greene, J.C. (1993). Data Analysis Strategies for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis, 15*(2), 195-207.
- Carney, K. J., & Stiefel, G. S. (2008). Long-term results of a problem-solving approach to response to intervention: Discussion and implications. *Learning Disabilities: A Contemporary Journal, 6*(2), 61-75.
- Chall, J. S., & Jacobs, V.A. (2003). Poor children's fourth-grade slump. *American Educator, 27*(1).
- Chall, J. S., Jacobs, V.A., & Baldwin, L.E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, Mass.: Harvard University Press.

- Chinn, C. A., Waggoner, M. A., Anderson, R. C., Schommer, M., & Wilkinson, I. A. (1993). Situated action in the small-group reading lesson: A microanalysis of oral reading error episodes. *American Educational Research Journal*, *30*, 361-392.
- Cho, J. (1998, April). *Rethinking curriculum implementation: Paradigms, models, and teachers' work*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Clark, C.M., & Peterson, P.L. (1986) Teachers' thought processes. In M.C. Wittrock, (Ed.), *Handbook of research on teaching* (3rd ed; pp. 255–296), New York: Macmillan.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, *41*, 778-820.
- Clotfelter, C. T., Ladd, H. F., & Vidgor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with fixed student effects. *Economics of Education Review*, *26*, 673-682.
- Cohen, J., Cohen, P., West, S., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cole, A.D. (2006). Scaffolding beginning readers: Micro and macro cues teachers use during student oral reading. *The Reading Teacher*, *59*(5), 450-459.
- Connolly, P. (2007). *Quantitative Data Analysis in Education: A Critical Introduction Using SPSS*. (pp. 95) New York: Routledge.
- Connor, C. M., Morrison, F. J., & Petrella, J. N. (2004). Effective reading comprehension

- instruction: Examining child by instruction interactions. *Journal of Educational Psychology*, 96(4), 682-698.
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research* (3rd Ed.; pp. 96-97). Los Angeles, CA: Sage
- Creswell, J.W. (1994). *Research design: Qualitative and quantitative approaches*. London: Sage.
- Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd, ed.). Upper Saddle River, NJ: Pearson.
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five traditions* (2nd Ed). Thousand Oaks, CA: Sage.
- Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209–240). Thousand Oaks, CA: Sage.
- Daly, E. J., Martens, B. K., Barnett, D., Witt, J. C., & Olson, S. C. (2007). Varying intervention delivery in response to intervention: Confronting and resolving challenges with measurement, instruction, and intensity. *School Psychology Review*, 36(4), 562–581.
- Dane, A. V., & Schneider, B. H. (1998) Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, 18, 23–45.
- Darling-Hammond, L. (2000). Teacher quality and student achievement. *Educational*

Policy Analysis Archives, 8(1). Retrieved from
<http://epaa.asu.edu/ojs/article/view/392>.

- Datnow, A., & Castellano, M. (2000). Teachers' responses to Success for All: How beliefs, experiences and adaptations shape implementation. *American Educational Research Journal*, 37, 775-799.
- Datta, L. (1997). A paradigmatic basis for mixed-method designs, in J. C. Greene and V. J. Caracelli (Eds.) *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms*. San Francisco, CA: Jossey-Bass.
- Deno, S. L., Fuchs, L. S., Marston, D. B., & Shin, J. (2001). Using curriculum-based measurement to develop growth standards for students with learning disabilities. *School Psychology Review*, 30, 507-524.
- Dickinson, D. K., & DeTemple, J. M. (1998). Putting parents in the picture: Maternal reports of preschoolers' literacy as a predictor of early reading. *Early Childhood Research Quarterly*, 13, 241-261.
- Dole, J. A., Nokes, J. D., & Drits, D. (2009). Cognitive strategy instruction. In G. G. Duffy & S. E. Israel (Eds.), *Handbook of research on reading comprehension* (pp 347-372). Hillsdale, NJ: Erlbaum.
- Duffy, G. G., & McIntyre, L. (1982). A naturalistic study of instructional assistance in primary grade reading. *Elementary School Journal*, 83, 15-23.
- Duffy, G. G., Roehler, L. R., Meloth, M. S., Vavrus, L. G., Book, C., Putnam, J., & Wesselman, R. (1986). The relationship between explicit verbal explanations during reading skill instruction and student awareness and achievement: A study of reading teacher effects. *Reading Research Quarterly*, 21(3), 237-252.

- Duffy, G. G., Roehler, L. R., & Rackliffe, G. (1986). How teachers' instructional talk influences students' understanding of lesson content. *The Elementary School Journal*, 87(1) 3-16.
- Duke, N.K. & Pearson, P.D. (2002). Effective Practices for Developing Reading Comprehension. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction (3rd ed)* (pp. 205-242) Newark, DE: International Reading Association.
- Eagan, B.A. (2009). Learning conversations and listening pedagogy: the relationship in student teachers' developing professional identities, *European Early Childhood Education Research Journal*, 17(1), 43–56.
- Ehri, L.C., Dreyer, L., Flugman, B., & Gross, A. I. (2007). Reading Rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal*, 44(2), 414–448.
- Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, 92(4), 605–619.
- Elmore, R., & Sykes, G. (1992). Curriculum policy. In P. Jackson (Ed.), *Handbook of research on curriculum* (pp. 185-215). New York: Macmillan.
- Emshoff, J. G., Blakely, C., Gottschalk, R., Mayer, J., Davidson, W. S., & Erickson, S. (1987). Innovation in education and criminal justice: Measuring fidelity of implementation and program effectiveness. *Educational Evaluation and Policy Analysis*, 9(4), 300-311.

- Englert, C. S., & Semmel, M. I. (1983). Spontaneous teacher decision making in interactive contexts. *Journal of Educational Research, 77*, 112-121.
- Englert, C. S., & Tarrant, K. L. (1995). Creating collaborative cultures for educational change. *Remedial and Special Education, 16*, 325-336.
- Finn, B., & Metcalfe, J. (2010). Scaffolding feedback to maximize long-term error correction. *Memory & Cognition, 38*(7), 951-961. doi:10.3758/MC.38.7.951
- Fisher, R. (2002). Shared thinking: Metacognitive modelling in the literacy hour. *Reading: Literacy and Language, 36*(2), 63-67.
- Flugum, K. R., & Reschly, D. J. (1994). Prereferral interventions: Quality indices and outcomes. *Journal of School Psychology, 32*, 1-14.
- Fryer, R.G. & Levitt, S.G. (2002). *Understanding the black-white test score gap in the first two years of school*. Cambridge, MA: National Bureau Of Economic Research. Retrieved from: <http://www.nber.org.proxy-um.researchport.umd.edu/papers/w8975>
- Fuchs, D., & Fuchs, L.S. (2009). Responsiveness-to-Instruction: Multi-level assessment and instruction as early identification and disability identification. *The Reading Teacher, 63*(1), 250-252.
- Fuchs, D., Fuchs, L.S., & Compton, D.L. (2004). Identifying reading disability by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly, 27*(4), 216-227.
- Fuchs, D., Fuchs, L. S., Mathes, P., & Simmons, D. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal, 34*, 174-206. <http://dx.doi.org/10.2307/1163346>

- Fuchs, D., Mock, D., Morgan, P., & Young, C. (2003). Responsiveness-to-intervention: Definitions, evidence and implications for the learning disabilities construct. *Learning Disabilities Research and Practice, 18*(3), 157-171.
- Fuchs, L. S. (n.d.). Project PROACT maze reading passages. Nashville, TN: Vanderbilt University.
- Fuchs, L. S., & Fuchs, D. (1990). Relations among the Reading Comprehension subtest of the Stanford Achievement Test and maze, recall, question answering, and fluency measures. Unpublished data.
- Fuchs, L. S., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. *School Psychology Review, 21*, 45-58.
- Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice, 13*, 204-219.
- Garrison, J. (1997). *Dewey and Eros: Wisdom and desire in the art of teaching*. New York: Teachers College Press.
- Gay, L. R., Mills, G. E., & Airasian, P. (2006). *Educational research: Competencies for analysis and applications*. (8th edition). NJ: Pearson.
- Gerber, M. M. (2005). Teachers are still the test: Limitations of response to instruction strategies for identifying children with learning disabilities. *Journal of Learning Disabilities, 38*, 516-524.
- Gersten, R., Baker, S. K., Haager, D., & Graves, A. W. (2005). Exploring the role of teacher quality in predicting reading outcomes for first-grade English learners: An observational study. *Remedial and Special Education, 26*(4), 197-206.

- Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools* (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from: <http://ies.ed.gov/ncee/wwc/publications/practiceguides/>
- Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C. & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, 71(2), 149-164.
- Gersten, R., Fuchs, L. S., Williams, J. R., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, 71, 279-321. <http://dx.doi.org/10.3102/00346543071002279>
- Glaser, B.G. (1965). The constant comparative method of qualitative analysis. *Social Problems*, 12(4), pp. 436-445.
- Glaser, B.G & Strauss, A.L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine Publishing Company.
- Glesne, C. (2006). *Becoming Qualitative Researchers: An Introduction*. (pp. 150) Boston: Pearson Education, Inc.
- Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22, 129-145.
- Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance of decision-

making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high stakes outcomes. *Scientific Studies of Reading*, 5(3), 257-288.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-on-one tutoring. *Applied Cognitive Psychology*, 9, 495-522.

Graham, S. & Harris, K.R. (2009) Almost 30 years of writing research: Making sense of it all with the wrath of Khan. *Learning Disabilities Research & Practice*, 24(2), 58–68.

Greene, J. C. (2001). Mixing social inquiry methodologies. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 251-258). Thousand Oaks, CA, Sage.

Gresham, F. M., Gansle, K. A., Noell, G. H., Cohen, S., & Rosenblum, S. (1993). Treatment integrity of school-based behavioral intervention studies: 1980–1990. *School Psychology Review*, 22, 254–272.

Gresham, F. M., MacMillan, D. L., Beebe-Frankenberger, M.E., & Bocian, K.M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disability Research & Practice*, 15, 198-205.

Hallahan, M. & Rosenthal, R. (1996). Statistical power: Concepts, procedures, and applications. *Behavioral Research Therapy*, 34(5/6), 489-499.

Hammill, D.D. (1990). On defining learning disabilities: An emerging consensus. *Journal of Learning Disabilities*, 23, 74-84.

Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). The Market for Teacher Quality. *National Bureau of Economic Research Working Paper* 11154.

- Harry, B., & Klingner, J. K. (2006). *Why are so many minority students in special education? Understanding race and disability in schools*. New York : Teachers College Press.
- Hattie, J.A., Gan, M. (2011). Instruction based on Feedback. In R. Mayer & P. Alexander, (Eds). *Handbook of Research on Learning and Instruction*. New York: Routledge.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1) pp. 81–112.
- Heck, R. H. (2007). Examining the relationship between teacher quality as an organizational property of schools and students' achievement and growth rates. *Educational Administration Quarterly*, 43(4), 399-432.
- Hoffman, J. V., O'Neal, S. F., Kastler, L. A., Clements, R. O., Segel, K. W., & Nash, M. F. (1984). Guided oral reading and miscue focused verbal feedback in second-grade classrooms. *Reading Research Quarterly*, 19(3), 367-384
- Hollenbeck, A. F. (2007). From IDEA to implementation: A discussion of foundational and future responsiveness-to-intervention research. *Learning Disabilities Research & Practice*, 22(2), 137–146
- Hsieh, H. & Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288.
- Huck, S. W. (2008). *Reading statistics and research*. (5th ed.) Boston: Pearson Education Inc.
- Individuals with Disabilities Education Improvement Act (IDEA) of 2004, PL 108–446, 20 USC §§ 1400 et seq.

- Jenkins, J. R., Heliotis, J. D., Stein, M. L., & Haynes, M. C. (1987). Improving reading comprehension by using paragraph restatements. *Exceptional Children, 54*(1), 54-59.
- Jenkins, J.R. & Jewell, M. (1993). Examining the validity of two measures of formative teaching: Reading aloud and maze. *Exceptional Children, 59*(5), 421–432
- Jenkins, J. R., & Leicester, N. (1992). Specialized instruction within general education: A case study of one elementary school. *Exceptional Children, 58*(6), 555-563.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33* (7) 14-26.
- Johnson, E., Mellard, D. F., Fuchs, D., & McKnight, M. A. (2006). *Responsiveness to intervention (RTI): How to do it*. Lawrence, KS: National Research Center on Learning Disabilities.
- Juel, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly, 31*, 268-289.
- Kavale, K. A. (2005). Effective intervention for students with specific learning disability: The nature of special education. *Learning Disabilities, 13*, 127-138.
- Kavale, K. A., & Reese, J. H. (1992). The character of learning disabilities: An Iowa profile. *Learning Disability Quarterly, 15*, 74–94.
- Kavale, K. A., Kauffman, J. M., Bachmeier, R. J., & LeFever, G. B. (2008). Response-to-intervention: Separating the rhetoric of self-congratulation from the reality of specific learning disability identification. *Learning Disability Quarterly, 31*(3), 135-150.
- Kendall, M. & Stuart, A. (1958). *The advanced theory of statistic*. New York: Hafner.

- Klingner, J. K., Cramer, E., & Harry, B. (2006). Challenges in the implementation of Success for All by four urban schools. *Elementary School Journal*, 106, 333-349.
- Klingner, J. K., & Edwards, P. (2006). Cultural considerations with response to intervention models. *Reading Research Quarterly*, 41, 108-117.
- Klingner, J. K., Vaughn, S., Hughes, M. T., & Arguelles, M. E. (1999). Sustaining research-based practices in reading: A three year follow-up. *Remedial and Special Education*, 20, 263-274.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.
- Kovaleski, J. F. (2007). Response to intervention: Considerations for research and systems change. *The School Psychology Review*, 36, 638-646.
- Kratochwill, T. R., Volpiansky, P., Clements, M., & Ball, C. (2007). Professional development implementing and sustaining multitier prevention models: Implications for response to intervention. *School Psychology Review*, 36, 618-631.
- Kreft, I.G.G. (1996). *Are multilevel techniques necessary? An overview, including simulation studies*. California State University, Los Angeles, CA.
- Landau, E. (1996). *Tropical Forest Mammals*. Danbury, CT: Children's Press.
- Lara, A. V., & Medley, D. M. (1987). Effective teacher behavior as a function of learner ability. *Professional School Psychology*, 2(1), 15-23.
- Levinson, S. S. C. (1983). *Pragmatics*. Cambridge, England: Cambridge University.
- MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). Gates-MacGinitie reading tests. Fourth edition. Forms S and T. Itasca, IL: Riverside.

- Manchester, K., Marcotte, A.M., & Matthews, W.J. (2004). *An initial reliability and concurrent validity investigation of the Dynamic Indicators of Basic Early Literacy Skills' Retell Fluency*. Paper presented at the meeting of the National Association of School Psychologists, Dallas, TX.
- Many, J. (2002). An exhibition and analysis of verbal tapestries: Understanding how scaffolding is woven into the fabric of instructional conversations. *Reading Research Quarterly, 37* (4), 376-407.
- Mariage, T. V. (1995). Why students learn: The nature of teacher talk during reading. *Learning Disability Quarterly, 18*(3) 214-234.
- Mariage, T. V., Englert, C. S., & Garmon, M. A. (2000). The teacher as "more knowledgeable other" in assisting literacy learning with special needs students. *Reading and Writing Quarterly, 16*, 299-336.
- Marston, D., Fuchs, L. S., & Deno, S. L. (1986). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. *Diagnostique, 11*(2), 77-90.
- Martin, J., Veldman, D. J., & Anderson, L. M. (1980). Within-class relationships between student achievement and teacher behaviors. *American Educational Research Journal, 17*(4), 479-490.
- Marzano, R. (2001). *A new era of school reform: Going where research takes us*. Aurora, CO: Mid-continent Research for Education and Learning.
- Mastropieri, M. A., & Scruggs, T. E. (2005). Feasibility and consequences of response to

- intervention: Examination of the issues and scientific evidence as a model for the identification of individuals with learning disabilities. *Journal of Learning Disabilities, 38*, 525-531.
- Mather, N., Hammill, D. D., Allen, E. A., & Roberts, R. (2004). *Test of Silent Word Reading Fluency*. Austin, TX: PRO-ED
- Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly, 40*, 148-182.
- McGill-Franzen, A. (2005). In the press to scale up, what is at risk? *Reading Research Quarterly, 40*(3), 366–370.
- McIntyre, E., Rightmyer E.C., & Petrosko J.P. (2008). Scripted and unscripted reading instructional models: Effects on the phonics and reading achievement of first grade struggling readers. *Reading & Writing Quarterly, 24*(4), 377-407.
- McKenzie, R. G. (2009). Obscuring vital distinctions: The oversimplification of learning disabilities within RTI. *Learning Disability Quarterly, 32* (4), 203-215.
- Mertzman, T. (2008). Individualising scaffolding: teachers' literacy interruptions of ethnic minority students and students from low socioeconomic backgrounds. *Journal of Research in Reading, 31*, 183–202. doi: 10.1111/j.1467-9817.2007.00356.x
- Meyer, L. A. (1986). Strategies for correcting students' wrong responses. *The Elementary School Journal, 87*(2), 227-241.
- Meyers, L.S., Gamst, G., & Guarino, A.J. (2006). *Applied Multivariate Research: Design*

- And Interpretation*. Thousand Oaks, CA: Sage Publications.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage.
- Morris, D. (2006). Using noncertified tutors to work with at-risk readers: an evidence-based model. *The Elementary School Journal*, 106(4), 351-362.
- Morrison, F.J., Bachman, H.J. & Connor, C.M. (2005). *Improving literacy in America: Guidelines from research*. New Haven, CT: Yale University Press.
- Myrberg, E., & Rosen, M., (2008). A path model with mediating factors of parents' education on students' reading achievement in seven countries. *Educational Research and Evaluation*, 14(6), 507–520
- National Institute of Child Health and Human Development. (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* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.
- Newman, A. P., Lewis, W., & Beverstock, C. (1993). *Prison literacy*. Philadelphia, PA: National Center on Adult Literacy.
- No Child Left Behind (NCLB) Act of 2001, 20 U.S.C.A. § 6301 *et seq.*
- Noell, G. H., Gresham, F. M., & Gansle, K. A. (2002). Does treatment integrity matter? A preliminary investigation of instructional implementation and mathematics performance. *Journal of Behavioral Education*, 11, 51-67.
- O'Brien, R.M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality and Quantity*, 41(5), 673-690.

- O'Donnell, C.L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 28*, 33-84.
- O'Sullivan, R. (2003). *Protecting sea turtles*. Washington, DC: National Geographic.
- Özerk, K. (2001). Teacher–student verbal interaction and questioning, class size and bilingual students' academic performance. *Scandinavian Journal of Educational Research, 45*(4), 353-367.
- Palincsar, A.S. (1986). Reciprocal teaching. *In Teaching reading as thinking*. Oak Brook, IL: North Central Regional Educational Laboratory.
- Partin, R.L. (1996). *Classroom Teacher's Survival Guide*. Englewood Cliffs, NJ: Prentice-Hall.
- Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 31*, 3-8. doi:10.1037/0278-7393.31.1.3
- Pentimonti, J., & Justice, L. (2010). Teachers' use of scaffolding strategies during read alouds in the preschool classroom. *Early Childhood Education Journal, 37* (4), 241-248.
- Penuel, W.R., & Means, B. (2004). Implementation variation and fidelity in an inquiry science program: Analysis of GLOBE data reporting patterns. *Journal of Research in Science Teaching, 41*, 294 – 315.
- Pianta, R. C., Belsky, J., Vandergrift, N., Houts, R. and Morrison, F. (2008). Classroom effects on children's achievement trajectories in elementary school. *American Educational Research Journal, 45*, 365-397.

- Pinnell, G. S., Lyons, C. A., Deford, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly, 29*(1), 8–39.
- Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking reliability and validity in content analysis. *Journal of Applied Communication Research, 27*, 258-284.
- Powell, K., & Kalina, C. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education, 130*(2), 241-250.
- Pressley, M., El-Dinary, P. B., Gaskins, I. W., Schuder, T., Bergman, J.L., Almasi, J., & Brown, R. (1992). Beyond direct explanation: Transactional instruction of reading comprehension strategies. *The Elementary School Journal, 92*(5), 513-555.
doi:10.1086/461705
- Pressley, M., & Wharton-McDonald R. (1997). Skilled comprehension and its development through instruction. *School Psychology Review, 26*, 448-466.
- Raphael, T. E., & Au, K. H. (2005). QAR: Enhancing comprehension and test taking across grades and content areas. *Reading Teacher, 59*, 206-221. <http://dx.doi.org/10.1598/RT.59.3.1>
- Reschly, D. J. (2003, December). *What if learning disabilities identification changed to reflect research findings?* Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention Symposium, Kansas City, MO.
- Reynolds C. R., & Shaywitz, S. E. (2009). Response to intervention: ready or not? Or from wait-to-fail to watch-them-fail. *School Psychology Quarterly, 24*(2) 130-145.

- Ritchey, K. D., Silverman, R. D., Montanaro, E. A., Speece, D. L., & Schatschneider, C. (2012). Effects of a tier 2 supplemental reading intervention for at-risk fourth-grade students. *Exceptional Children, 78*(3), 318-334.
- Ritchey, K. D., Speece, D. L., Silverman, R.D., & Montanaro, E.A. (n.d.). Assessment of strategy knowledge and use for information text (ASKIT): Sea Turtles (Form A) and Strange Animals (Form B). University of Maryland Reading Research Projects. College Park, MD: University of Maryland.
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review, 94*(2), 247-252.
- Rodgers, E.M. (1999). Understanding teacher and student talk during literacy instruction in a one-to-one tutoring setting. (Doctoral Dissertation, The Ohio State University). Retrieved from:
http://etd.ohiolink.edu/view.cgi?acc_num=osu1242845174
- Rodgers, E.M. (2004). Interactions that scaffold reading performance. *Journal of Literacy Research, 36*(4), 501-532 doi: 10.1207/s15548430jlr3604_4
- Rossmann, G.B., & Wilson, B.L. (1994). Numbers and words revisited: Being shamelessly eclectic. *Quality and Quantity, 28*(3), 315-327.
- Ruiz-Primo, M. A. (2006). *A multi-method and multi-source approach for studying fidelity of implementation*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Saldana, J. (2009). *The coding manual for qualitative researchers*. Los Angeles, CA: SAGE.
- Sammons, P., Day, C., Kington, A., Gu, Q., Stobart, G., & Smees, R. (2007). Exploring

- variations in teachers' work lives and their effects on pupils: Key findings and implications from a longitudinal mixed methods study. *British Education Research Journal*, 33, 681-701.
- Samph, T. (1974). Teacher behavior and the reading performance of below-average achievers. *The Journal of Educational Research*, 67(6), 268-270.
- Sanacore, J. & Palumbo, A. (2009). Understanding the fourth-grade slump: Our point of view. *The Educational Forum*, 73, 67-74.
- Savage, R. Carless, S., & Erten, O. (2009). The longer-term effects of reading interventions delivered by experienced teaching assistants. *Support for Learning*, 24(2), 95-100.
- Sayre, A.P. (2003). *Tropical Rainforests*. New York, NY: Scholastic Reference.
- Schlechter, M. J., & Milevsky, A. (2010). Parental level of education: Associations with psychological well-being, academic achievement, and reasons for pursuing higher education in adolescence. *Educational Psychology*, 30, 1-10.
- Scrimsher, S., & Tudge, J. (2003). The teaching/learning relationship in the first years of school: Some revolutionary implications of Vygotsky's theory. *Early Education and Development*, 14, 293-312.
- Seltman, H.J. (2012). *Experimental Design and Analysis*. Retrieved from:
http://www.stat.cmu.edu/_hseltman/309/Book/Book.pdf
- Semmel, D. S. (1977, April). *The effects of teacher training of decision-making*. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Shaywitz, S. E., Escobar M. D., Shaywitz B. A., Fletcher J. M., & Makuch R. (1992).

- Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, 326(3), 145-150.
- Shin, J., Deno, S.L., & Espin, C.A. (2000). Technical adequacy of the maze task for Curriculum-Based Measurement of reading growth. *Journal of Special Education*, 34, 164-172.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153 -189. doi:10.3102/0034654307313795
- Slavin, R. E., Lake, C., Davis, S., & Madden, N. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6, 1-26.
- Smith, T. W., Snyder, C. R., & Handelsman, M. M. (1982). On the self-serving function of an academic wooden leg: Test anxiety as a self-handicapping strategy. *Journal of Personality and Social Psychology*, 42, 314–321.
- Snow, C. (2002). Reading for Understanding: Toward an R&D Program in Reading Comprehension. Santa Monica, CA: RAND Corporation. Retrieved from http://www.rand.org/pubs/monograph_reports/MR1465.
- Spear-Swerling, L. (2009). A literacy tutoring experience for prospective special educators and struggling second graders. *Journal of Learning Disabilities*, 42(5), 431-443.
- Speece, D. L., & Case, L. P. (2001). Classification in context: An alternative approach to identifying early reading disability. *Journal of Educational Psychology*, 93, 735-749.
- Speece, D. L., Ritchey, K. D., Silverman, R., Schatschneider, C., Walker, C. Y., &

- Andrusik, K. N. (2010). Identifying children in middle childhood who are at risk for reading problems. *School Psychology Review, 39*(2), 258-276.
- Spera, C. (2006). Adolescents' perceptions of parental goals, practices, and styles in relation to their motivation and achievement. *Journal of Early Adolescence, 26*, 456-490.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.
- Stevenson, H. W. and Cruse, D. B., (1961). The effectiveness of social reinforcement with normal and feeble-minded children. *Journal of Personality, 29*, 124–135.
- Strauss, A.C. & Corbin, J.M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA: Sage Publications.
- Suhr, K.A., Hernandez, D.A., Grimes, D., & Warschauer, M. (2010). Laptops and fourth-grade literacy: Assisting the jump over the fourth-grade slump. *Journal of Technology, Learning, and Assessment, 9*(5). 5-45.
- Sum, A., Kirsch, I. & Taggart, R. (2002). *The Twin Challenges of Mediocrity and Inequality: Literacy in the U.S. from an International Perspective*. Princeton, NJ: Educational Testing Service.
- Svensson, I., Lundberg I., & Jacobson, C. (2003) The nature of reading difficulties among inmates in juvenile institutions. *Reading and Writing 16*(7), 667-691, doi: 10.1023/A:1025832815286
- Tashakkori, A. & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA, Sage Publications.
- Taylor, B. M., Pearson, P. D., Clark, K., & Walpole, S. (2000). *Effective schools and*

- accomplished teachers: Lessons about primary grade reading instruction in low-income schools. *Elementary School Journal*, 101(2), 121-166.
- Taylor, B. M., Pearson, P. D., Peterson, D. S., & Rodriguez, M. C. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *Elementary School Journal*, 104, 3-28.
- Taylor, B. M., Peterson, D. P., Pearson, P. D., & Rodriguez, M. C. (2005). The CIERA school change framework: An evidence-based approach to professional development and school reading improvement. *Reading Research Quarterly*, 40(1), 40-69.
- Tashakkori, A. & Teddlie, C. (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.
- Tharp, R.G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. Cambridge: Cambridge University Press.
- Thompson, B., Diamond, K. E., McWilliam, R. Snyder, P., & Snyder, S. W. (2005). Evaluating the quality of evidence from correlational research for evidence-based practice. *Exceptional Children*, 71(2), 181-194.
- Thompson, T. (1999). *Underachieving to protect self-worth: Theory, research and interventions*. Aldershot: Ashgate.
- Thompson, T., & Richardson, A. (2001). Self-handicapping status, claimed self-handicaps and reduced practice effort following success and failure feedback. *British Journal of Educational Psychology*, 71, 151-170.
- Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Lindamood, P., Rose, E., Conway, T., &

- Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91*, 579-593.
- Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K., Conway, T., & Rose, E. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34*, 33-58.
- U.S. Department of Education. Office of Educational Research and Improvement. National Center for Education Statistics. (2000). NAEP 1999 Trends in Academic Progress: Three Decades of Student Performance, NCES 2000-469, by J.R. Campbell, C.M. Hombro, and J. Mazzeo. Washington, DC: 2000.
- U.S. Department of Education. (2009). Race to the Top executive summary. Retrieved from <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf>
- Vaughn, S., & Fuchs, L. S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice, 18*(3), 137-146.
- Vaughn, S., Fletcher, J.M., Francis, D.J., Denton, C.A., Wanzek, J., Wexler, J., Cirino, P.T., Barth, A.E., & Romain, M.A. (2008). Response to intervention with older students with reading difficulties. *Learning and Individual Differences, 18*(3), 338-345. doi: 10.1016/j.lindif.2008.05.001
- Vaughn, S., Denton, C. A., & Fletcher, J. M. (2010). Why intensive interventions are necessary for students with severe reading difficulties. *Psychology in the Schools, 47*(5), 432-444.

- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M.B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88*, 601-638.
- Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities, 39*(2), 157-169.
- Vygotsky, L. S. (1987). *The collected works of L. S. Vygotsky, Vol. 1: Problems of general psychology*. New York: Plenum Press.
- Vygotsky, L.S. (1978). Interaction between Learning and Development (pp. 79-91). *In Mind in Society*. Cambridge, MA: Harvard University Press.
- Wanzek, J., & Vaughn S. (2007). Research-based implications from extensive early reading interventions. *The School Psychology Review, 36*, 541-561.
- Wanzek, J., & Vaughn, S. (2008). Response to varying amounts of time in reading intervention for students demonstrating insufficient response to intervention. *Journal of Learning Disabilities, 41*, 126–142.
- Wharton-MacDonald, R., Pressley, M., & Hampston, J. M. (1998). Literacy instruction in nine first-grade classrooms: Teacher characteristics and student achievement. *The Elementary School Journal, 99*, 101–128.
- Wiberg, M. & Sundström, A. (2009). A comparison of two approaches to correction of

- restriction of range in correlation analysis. *Practical Assessment, Research & Evaluation, 14* (5). Available online: <http://pareonline.net/getvn.asp?v=14&n=5>
- Willie, C. V. (2001). The contextual effects of socioeconomic status on student achievement test scores by race. *Urban Education, 36*(4), 461.
- Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*, 89-100
- Wood, B.K., Umbreit, J., Liaupsin, C.J., & Gresham, F.M. (2007). A treatment integrity analysis of function-based intervention. *Education and Treatment of Children, 30*, 105-120.
- Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. In B. Wildemuth (Ed.), *Applications of Social Research Methods to Questions in Information and Library Sciences*. Englewood, Colorado: Libraries Unlimited.