MULTIDIMENSIONAL LITERACY DEVELOPMENT:

EXAMINING CHILD ENGAGEMENT AND TEACHER INSTRUCTION

Cathleen Carolyn Reynolds

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education (School Psychology).

Chapel Hill
2014

Approved by:

Sandra B. Evarrs

Leigh A. Hall

Steven E. Knotek

Rune J. Simeonsson

Catherine Zimmer

ABSTRACT

Cathleen Carolyn Reynolds: Multidimensional Literacy Development:
Examining Child Engagement and Teacher Instruction
(Under the direction of Rune J. Simeonsson)

Given the importance of federal legislation in combating educational inequities, particularly those in reading development, the identification of specific school-based variables that contribute to reading proficiency for struggling readers is paramount. While federal education policy, such as the No Child Left Behind (NCLB) Act of 2001 (2002) focuses primarily on direct instruction of specific literacy skills and characteristics of teacher quality, the literature does not show strong support for isolating these variables to improve children's reading skills in public schools. A growing body of literature has focused on the connection between child classroom behavioral engagement and reading achievement. This study utilizes data from the Early Childhood Longitudinal Study - kindergarten class (ECLS-K) to examine of child behavioral engagement skills and teacher instruction in relation to children's literacy development in public elementary schools. Specifically, three research questions were examined using multivariate regression analyses. First, the contribution of child behavioral engagement skills to child literacy outcomes in third and fifth grades were examined above and beyond the contribution of their kindergarten literacy skills and participation in Reading First interventions.

Second, the contribution of teacher's instructional choices on child literacy outcomes in third and fifth grades were examined above and beyond the contribution of teachers' Highly Qualified Teacher (HQT) status and children's participation in additional instruction time. Finally, within-year analyses examined individual and interaction effects of child behavioral engagement and teacher instructional choices on within-year child literacy scores. Support was found for the first hypothesis, but no support was found for the second hypothesis. Further, instructional time, teacher certification, and level of teacher education did not make any significant contribution to child reading scores. This research clarifies and extends current knowledge of constructs highlighted in NCLB and child classroom behavioral engagement. Findings of this research, in conjunction with future efforts, offer information for policymakers to develop evidence-based educational policies and improve the quality of public education for our children.

To my daughters, Emma and Margaret.
You can achieve anything.

To my husband, Pete.
Your unconditional support, encouragement, and love made this possible.

And, to my parents,
for always encouraging me to do my best.

ACKNOWLEDGEMENTS

I would like to acknowledge the contributions of each member of my committee. Your thoughtful suggestions and edits improved the quality of this project and helped me to improve my writing skills. I am truly grateful for each of your input. I would also like to specifically thank my advisor and chair of my dissertation committee, Rune Simeonsson. Throughout my graduate school experiences in coursework and research efforts, my greatest influence in conceptualizing the developing child has come from working with Rune. I am so grateful to have been able to get a glimpse of the important work that Rune conducts in the name of child health and positive child development, worldwide. I will remain committed to prevention and early intervention efforts throughout my career because of this influence. I would also like to specifically thank Cathy Zimmer. I met Cathy through my work on this project, and her input on the data preparation and statistical analysis was invaluable. I truly believe that I could not have completed this project without her specialized expertise and encouragement. I hope to remain in touch with both Rune and Cathy as I begin my career. I would also like to acknowledge the support and encouragement from my friends and family. The quiet support and encouragement of my parents, sister, brother, and sister-in-law was always present in my mind. My extended Reynolds family, who expressed nothing but wonderful support and encouragement along the way - thank you. And to numerous friends who cheered me on, most especially Erinne Selim, whom I have known my whole life, yet never ceases to amaze me with her thoughtfulness and

expressions of kindness. Finally, I must acknowledge the significant sacrifice of my husband and two daughters during this long process - both in the loss of my presence and the loss of their own personal time. I am so very lucky to have you in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACYP	Administration on Children, Youth, and Families
ECLS-K	Early Childhood Longitudinal Study – Kindergarten Class of 1998-1999
HQT	Highly Qualified Teacher, as defined in NCLB
IRT	Item Response Theory
MBA	Woodcock-McGraw-Werder Mini Battery of Achievement
MCMC	Bayesian iterative Markov chain Monte Carlo
MI	Multiple Imputation
MVN	Multivariate Normal Multiple Imputation
NAEP	National Assessment for Educational Progress
NAEYC	National Association for the Education of Young Children
NASP	National Association of School Psychologists
NCES	National Center for Education Statistics
NCLB	The No Child Left Behind Act of 2001
NGACBP	National Governors Association Center for Best Practices & Council of Chief State School Officers
NICHD	National Institute of Child Health and Human Development
OSEP	Office of Special Education Programs
PSU	Primary Sampling Unit
REACT	Responsive Environmental Assessment for Classroom Teaching
RTI	Response to Intervention
RVI	Relative Increase in Variance

SES	Socioeconomic status
SPSS	Statistical Package for the Social Sciences
SRS	Social Rating Scale
SSRS	Social Skill Rating System
T-SRS	Teacher Social Rating Scale
US	United States
VIF	Variance Inflation Factor
ZPD	Zone of Proximal Development

LIST OF SYMBOLS

β	Beta
CI	Confidence Interval
F	F-Score
M	Imputed dataset
μ	Mathematical mu representing the mean or average score
p	p-value
R^2	R-squared
SD	Standard deviation
t	t-score

CHAPTER I: INTRODUCTION

Two primary goals of NCLB (2002) are to improve the academic achievement of underserved populations and for all students to be proficient in reading by the 2013-14 school year. This ambitious goal is addressed through a broad multifaceted intervention approach within the Act, including two major areas of focus for elementary school aged intervention: children's early literacy skills and teacher quality. Specifically, Reading First interventions and HQT status are highlighted in NCLB as essential components of a high quality education.

This research examined the impact of Reading First and HQT on the literacy skills of struggling readers in public elementary schools in concert with additional research suggesting that child behavioral engagement and teacher instructional choices can contribute to child literacy outcomes in public schools. Specifically, the contribution of child behavioral engagement skills to third and fifth grade child literacy outcomes were examined in relation to the contribution of children's kindergarten literacy skills and their participation in Reading First interventions. Second, the impact of teacher's instructional choices on child third and fifth grade literacy outcomes were examined in relation to the contribution of teachers' HQT status and children's participation in increased time on reading instruction. Finally, multiple regression analyses were utilized to determine the individual and interaction effects of child behavioral engagement and teacher instructional choices on within-year child literacy scores.

Evaluating the differential contribution of components of NCLB to child literacy outcome scores is necessary to ensure that federal education policy and funding are focused on empirically-validated facets of education. Although NCLB regulations have been in place since 2002, the National Assessment for Educational Progress (NAEP) reported in 2009 that 65% of fourth graders in the United States (US) were not proficient in grade-level reading skills (National Center for Education Statistics, NCES, 2014). Further, research in this area has demonstrated that children who are struggling with reading achievement in early elementary school tend to be members of underserved populations, including minority (Snow, Burns, & Griffin, 1998; West, Denton, & Germino-Hausken, 2000) and low socioeconomic status (SES; Denton & West, 2002; Snow et al., 1998) that are often referred to as "at risk" populations (Coie et al., 1993). These populations of "at risk" children are precisely those targeted through NCLB legislation.

Some previous research suggests that improvements in achievement were found following direct instruction in early literacy skills (i.e., Kaverz, 2002; National Institute of Child Health and Human Development; NICHD, 2000; Snow et al., 1998) and learning from a teacher with full certification (Chatterji, 2006; Goldhaber & Brewer, 1997; Goldhaber & Brewer, 2000; Hawk, Coble, & Swanson, 1985; Kukla-Acevedo, 2009; Nye, Konstantopoulos, & Hedges, 2004; Wayne & Youngs, 2003). Yet, after years of implementation in public schools, the achievement gap remains (NAEP, 2014). A growing body of literature has begun to report on additional variables that may improve the effectiveness of interventions funded through NCLB. A child's level of engagement with academic materials and a teacher's instructional choices are two variables that have been associated with improved achievement (Block, Parris, Reed,

Whiteley, & Cleveland, 2009; Mashburn et al., 2008; McClelland, Acock, & Morrison, 2006; Parlardy & Rumberger, 2008; Schaefer & McDermott, 1999; Spira, Bracken, & Fischel, 2005; Yen, Konold, & McDermott, 2004). Yet, despite the strong support linking both a child's behavioral engagement and a teacher's instructional style with child achievement outcomes. Current literature is in need of definitive research examining literacy outcomes with a population of children who begin school lacking essential pre-literacy skill development. It is these children, primarily targeted by NCLB, who require greater achievement gains within the same timeframe as their better-prepared peers in order to reach grade-level proficiency in reading.

Additionally, few results have been reported in the literature regarding the impact of children's behavioral engagement on their achievement outcomes in the later elementary school years (Alexander, Entwisle, & Dauber, 1993; McClelland et al., 2007; McClelland Morrison, & Holmes, 2000; Ponitz, Rimm-Kaufman, Grimm, & Curby, 2009), and inconsistent results have been reported regarding the impact of kindergarten behavioral engagement on reading outcomes in later elementary school years (McClelland et al., 2006). Furthermore, studies addressing the impact of kindergarten behavioral engagement on reading outcomes have largely focused on within-year results, rather than examining longitudinal data covering the entirety of elementary school (Alexander et al., 1993; McClelland et al., 2007; Ponitz et al., 2009). Finally, to date, no study has examined the potential differential impact of behavioral engagement skills measured during later elementary school years, nor has any study in this area of research included data regarding the potential impact of a child's participation in Reading First interventions.

With regard to teacher instructional choices, few studies have reported results that were specific to reading achievement (Mashburn et al., 2008), and previous results tend to be based on samples that are generalizable to an overall population of elementary students with diverse learning levels (Block et al., 2009; Mashburn et al., 2008), rather than specifically examining the impact of teacher instructional choices on the achievement of students who are already struggling or at-risk for poor achievement (Coie et al., 1993). Of the studies that did focus results on at-risk populations inconsistent results were seen (Hamre & Pianta, 2005; Xue & Meisels, 2004). A longitudinal study focusing on a population of struggling early readers from kindergarten through fifth grade is needed to clarify these results.

The longitudinal design of ECLS-K provides a unique opportunity to examine both child and teacher variables within a nationally representative sample of children who began school struggling to achieve essential pre-literacy skills in public school. This design clarifies previous literature and demonstrates whether a unique predictive effect of child behavioral engagement and teacher instructional choice exists for these struggling readers at different points across elementary school. Child behavioral engagement, teacher instructional style, and teacher qualifications were measured through the use of teacher report and examined with respect to their contribution to child reading achievement. Additional variables including race, SES, and child disability status were measured through parent report. Hierarchical regressions were used to examine both the child and teacher research questions, using variables from kindergarten through the fifth grade to predict third and fifth grade reading outcomes. Finally, multiple regression was used for within-year analyses on kindergarten, first, third, and fifth grades.

This research is intended to be a comprehensive examination of the impact of both child behavioral engagement and teacher instructional style on struggling readers' achievement over the course of elementary school, accounting for children's early reading status, participation in Reading First interventions, and characteristics of teacher quality. This study fills gaps in the current literature by examining these questions using a longitudinal, nationally representative sample of children who were performing poorly in reading achievement at the end of kindergarten.

In Chapter 2, literature describing struggling readers, the achievement gap, Reading First, HQT, child behavioral engagement skills, and teacher instructional choices will be reviewed and discussed in relation to the literacy achievement of struggling readers. In Chapter 3, research hypotheses, procedures and data preparation will be reviewed in detail. In Chapter 4, results of each research question will be reviewed and displayed in figures and tables at the end of this document. Finally, in Chapter 5, results of each research question will be discussed in relation to what was known from previous literature and NCLB. Conclusions, limitations and future directions of this area of research are discussed.

CHAPTER II: LITERATURE REVIEW

To define the scope of this research, this review of the literature begins with an examination of the current status of reading skills for public elementary schoolchildren in the United States. The achievement gap, long-term individual and educational correlates of academic reading development, and the prediction of reading scores from early indicators are reviewed. Federal policy addressing reading intervention is examined in concert with the bio-ecological theory of child development to provide a conceptual framework for examining child and teacher effects on reading development. Literature examining the impact of child behavioral engagement and teacher instructional style on child reading achievement discussed to provide the foundation for the primary research questions.

Childhood Reading

The ability to read and comprehend text is an essential component of success in today's society. Yet, in 2009, the NAEP reported that 65% of fourth grade children scored below grade-level proficiency on standardized tests of reading skills (NCES, 2014). The issue of reading development has been widely researched, including the relative stability and longstanding effects of reading scores (Butler, Marsh, Sheppard & Sheppard, 1985; Chatterji, 2006; Hanson & Farrell, 1995; Juel 1988; NCES, 2014; Stanovich, 1986), and the prediction of scores from early reading skills, such as phonological awareness and letter knowledge (Bishop, 2003; Duncan et al., 2007; Kirby, Parrilla, & Pfeiffer, 2003; Lonigan, Burgess, & Anthony, 2000; O'Connor &

Jenkins, 1999; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004). The results of previous work produced a deeper understanding of childhood reading development and ultimately contributed to improvements in reading intervention services for struggling students. Yet, it is clear that in order to improve the odds for young children who are struggling to achieve proficiency in reading, more work is needed to clarify and improve upon current knowledge.

Struggling readers. Before discussing factors that may contribute to the achievement of children who are below proficiency in reading, it is important to understand how this population of children has been conceptualized in previous literature. Struggling readers have been operationalized and described in the literature in a number of ways, covering both child background characteristics, performance on classroom assessments, and/or child scores on standardized and non-standardized tests of reading skills. Further, individual child characteristics, such as race (Snow et al., 1998; West et al., 2000; Wiggan, 2007), SES (Denton & West, 2002; Eamon, 2002; Singh, 2012; Sirin, 2005; Snow et al., 1998), age at kindergarten entry (Justice, Invernizzi, Geller, Sullivan, & Welsch, 2005), and gender (Denton & West, 2002; Justice et al., 2005; Phillips, Norris, Osmond, & Maynard, 2002) have been repeatedly linked in previous literature with the development of early reading skills and discussed as potential areas of risk (Coie et al., 1993).

Researchers posit that children who are at-risk (e. g., low SES) often are found to have multiple characteristics of risk, thus decreasing the odds of long-term educational success (Downer, Rimm-Kaufman, & Pianta, 2007; Rathbun, West, & Walson, 2005; Vitaro, Brendgen, Larose, & Tremblay, 2005). Further, children with an accumulation of risk factors tend to demonstrate poorer reading skills on school assessments (Livingston & Wirt, 2004; Zill & West,

2001) and show a greater risk of long-term negative secondary outcomes (Rathbun et al., 2005; Smokowski, Mann, Reynolds, & Fraser, 2004), including grade retention (Baydar, Brooks-Gunn, & Furstenberg, 1993), school failure (Mather & Ofiesh, 2005), and/or school dropout (Cairns, Cairns, & Neckerman, 1989; Korhonen, Linnanmaki, & Aunio, 2014; Vitaro et al., 2005). In fact, previous findings from ECLS-K demonstrated an 11-point difference between the kindergarten through third grade reading gains of children with no risk factors versus those with two or more risk factors (Livingston & Wirt, 2004). Indeed, it is well understood within the literature that children who are identified as struggling readers early in formal schooling are likely to also hold characteristics of risk.

While struggling readers are often found to hold multiple characteristics of risk, they are primarily identified in research through performance on reading assessments, or on an end-of-year high stakes assessment (Melekoglu, 2011; Reutzel, Petscher, & Spichtig, 2012; Roberts, Vaughn, Fletcher, Stuebing, & Barth, 2013). Indeed, this may be appropriate as current trends in public education have moved toward the use of high stakes testing in schools to determine academic proficiency and make grade-promotion decisions (NCLB, 2002). Each of these methods of identification use a cutoff point of reading performance on a one-time test.

Others have defined struggling readers through repeated examples of their in-school achievement in reading such as in-school grades, and in-class, teacher-administered 1:1 assessments of specific reading skills, repeated over time, to rank whether children were profiting from reading instruction and whether they were performing at achievement levels that were considered below, at, or above grade level (Vernon-Feagans et al., 2010). These tests, often termed Progress Monitoring within a Response to Intervention (RTI) framework, are commonly

thought to be accurate measures of a child's in-class achievement, measured over time to demonstrate whether children are making necessary improvements in skills as they are exposed to the reading curricula in class.

Due to the complexity of data-based decision making such as what is seen within a RTI framework, and the complexity of data collection within ECLS-K , this level of identification was not possible in the current research. Indeed, for the purposes of the current research, struggling readers were conceptualized as those who performed below a cutoff score on a one-time test capturing a snapshot of children's literacy skills as they completed kindergarten. Children who scored below the cutoff score were determined to be struggling with early reading skills. To further describe this sample of struggling readers, their proficiency on specific skills at the time of the test and multiple indicators of risk were explored and discussed.

Achievement gap. A child's individual characteristics and background, such as SES, in combination with child achievement or skill development present a multifaceted approach to describing children who are struggling. Research has shown that early childhood experiences are a strong contributor to the skill development and background knowledge of first time kindergarteners (Walberg & Tsai, 1983; West et al., 2000; Zill & West, 2001), and when children are asked to meet the educational and behavioral demands of the classroom context, children naturally draw from their early experiences for explanation (Rimm-Kaufman, Pianta, & Cox, 2000; Whitehurst & Lonigan, 1998). Those who lack a strong foundation of early academic and behavioral experiences must make greater learning gains than their better-prepared peers in order to reach equivalent achievement status as their better-performing peers by the end of each school year (Bast & Reitsma, 1997; Juel, 1988; Stanovich, 1986; Walberg &Tsai, 1983). As a result,

empirically-validated interventions to assist in improving academic and behavioral skills for children who present with multiple characteristics of risk, particularly those who present with poor early achievement during early elementary school are essential to reach the ambitious goal of reading proficiency for all children.

Without appropriate empirically-validated intervention, children who begin school struggling with reading skills are far less likely to reach grade-level expectations. Matthew effects (i.e., the widening gap, over time between children with strong academic reading skills and those who require more assistance in developing their academic reading skills; Stanovich, 1986) in reading skill development pose a considerable challenge for successful outcomes. The Matthew effect phenomenon and the relative stability of children's reading skills (Butler et al., 1985; Juel, 1988; Stanovich, 1986; Walberg & Tsai, 1983) have inspired a wealth of research examining the long-term predictive relationship between early reading skills (e.g., phonological awareness), long-term school achievement (Cunningham & Stanovich, 1997; Juel, 1988; McClelland et al., 2006; McClelland et al, 2000; NICHD, 2000; Scarborough & Parker, 2003; Torgeson et al., 1999), and adult contributions to society (National Association for the Education of Young Children, NAEYC, 1998). Yet others have reported results refuting evidence of the stability of reading during elementary school (Phillips et al., 2002) and of Matthew effects occurring within the school year (McCoach, O'Connell, Reis, & Levitt, 2006), suggesting that the relationship between early reading skills and later achievement may be mediated by other variables.

While school reading interventions focused primarily on the targeted training of specific early literacy skills (e.g., phonological awareness) have achieved some success in improving children's reading achievement (Elbro & Petersen, 2004; Lane, O'Shaughnessy, Lambros, Gresham, & Beebe-Frankenberger, 2001; Otaiba et al., 2008), children from at-risk backgrounds continue to achieve below their peers from more stable backgrounds (Elbro & Petersen, 2004; Lane et al., 2001; Otaiba et al., 2008). However, multivariate early childhood interventions such as the Abecedarian (Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002) and Perry Preschool programs (Schweinhart et al., 2005) using specialized curricula designed to target academic goals have reported positive results into adulthood. Current trends in the literature highlight not only the impact of specific reading skills, but also the impact of empirically-validated multivariate interventions that better account for the complex interplay of variables contributing to reading development.

Conceptual Framework

Though the complexity of reading development is not fully understood, this research narrowed the lens to focus specifically on influences of childhood reading development occurring within the child's immediate classroom environment across years of elementary school using a framework that highlights not only *what* early skills have developed, but also *how* this development occurs. The bio-ecological model posits that development is a dynamic interactive process occurring within multiple immediate and distal environments of the child over time (Bronfenbrenner & Morris, 1998). Within the bio-ecological model, proximal processes are regular and increasingly complex interactions between the developing child and persons (e.g., teachers, peers) or objects (e.g., instructional materials) within the child's environment (e.g.,

classroom) over time (e.g., a school year). According to the bio-ecological theory, a child's reading development will vary uniquely based on the quality of the proximal processes or interactions within the child's environment. Although the characteristics of both the developing child and persons within the child's environment make important contributions to developmental outcomes, proximal processes are established within the model as the "engines of development" (Bronfenbrenner & Morris, 1998, p. 999).

> The power of such processes to influence development is presumed, and shown, to vary substantially as a function of the characteristics of the developing *Person*, of the immediate and more remote environmental *Contexts*, and the *Time* periods, in which the proximal processes take place (Bronfenbrenner & Morris, 1998, p. 994).

Bronfenbrenner and Morris (1998) further posited that positive proximal processes impact achievement outcomes by leading to greater developmental gains, and by protecting "against effects of disadvantaged and disruptive environments" (p. 1004). This premise suggests that interventions focused on strengthening *proximal processes* in the classroom environment, and examining how the interactions with these variables might change over time, could produce robust results in combating poor reading achievement.

Yet, federal legislation aimed at reducing the achievement gap did not address proximal processes in the classroom. Portions of NCLB that pertained to elementary school-aged reading development addressed specific instructional and teacher characteristics; the Act failed to address complex processes, such as the child's level of engagement with educational materials and the effect of teachers' instructional style, over time.

The major component of NCLB targeting reading skills, the Reading First program (NCLB, 20 U.S.C. §§ 6361-6368 [2002]), is another example of this discrepancy. Although the program highlighted five essential components of reading (phonemic awareness, phonics, vocabulary, fluency, and comprehension) as the primary focus of funded interventions, the legislation was absent of any mention of proximal processes, such as a teacher's instructional style or a child's engagement with the academic materials. While the five essential components in NCLB come from a strong empirical foundation derived from the literature at the time of the 2002 reauthorization (i.e., Kaverz, 2002; NICHD, 2000; Snow et al., 1998), it is clear that focusing interventions primarily on these five academic components has not met the goal of making all children proficient readers by the third grade (Gamse et al., 2008; NCES, 2014).

Specifically, the Reading First Impact Study reported a significant increase in the amount of first and second grade classroom instructional time spent on the five early reading skills highlighted by the program, but found no significant impact on children's fourth grade reading comprehension scores or on their engagement with print (Gamse et al., 2008). Additionally, the Nation's Report Card reported that only 65% of fourth grade public school children achieved at or above proficiency in reading (NCES, 2014). In other words, although children are spending more instructional time on specific reading skills and have improved their standardized test scores, there appears to be no subsequent improvement in reading comprehension, children are not spending more time engaged in reading, and 65% of children are not proficient in reading by the end of fourth grade (NCES, 2014). While it is true that empirically supported interventions (addressing the five essential components of reading) funded by Reading First may have contributed to improvements in children's standardized test scores, a more multifaceted approach

to intervention, particularly one addressing proximal processes and the changing effects of the processes over time, may better support young struggling readers.

Likewise, NCLB focused on identifying HQT using qualifications such as demonstration of knowledge in subject areas (typically through a passing score on a state certification test) and certification status (NCLB, 20 U.S.C. § 7801 [2002]), rather than considering proximal processes, such as how a teacher presents academic materials to students. Further, although context-specific variables such as duration of instructional time and type of instructional materials were highlighted in the Act, the processes by which a child learns were largely omitted. This type of approach, emphasizing one learning domain (e.g., early academic skills) over another (e.g., behavioral engagement), is unsupported by research as a method of developing successful interventions (Scott-Little, Kagan, & Frelow, 2006). However, no studies, to date, have examined whether the inclusion of an alternate learning domain, such as learning-related behavioral development, might strengthen interventions supported by NCLB.

Although the bio-ecological theory suggests that proximal processes significantly impact learning and development, these processes have not yet been adequately addressed within major federal education policy. To date, no literature has examined the effect of process variables on reading outcomes relative to variables included in NCLB (2002). Proximal processes surrounding elementary school reading development in a classroom environment should be further explored in research efforts.

Proximal Processes in Classroom Reading Instruction

When examining discrepancies in reading instruction with the goal of reducing achievement disparities, it is wise to look not only at aspects of the academic curricula that are associated with later achievement (e.g., phonological awareness), but also at the interaction of proximal processes occurring, in context, between persons (or between persons and objects) over time. During elementary school reading instruction, one must consider not just *what* specific academic skills are presented, but also consider *how* they are presented and *how* individual children are receiving the instruction. Learning and development are posited to be bi-directional and interactive (Bronfenbrenner & Morris, 1998). This research examined whether specific proximal processes (i.e., a child's behavioral presentation and a teacher's instructional presentation) contributed to reading outcome scores above and beyond the contribution of the child and teacher characteristics that are currently identified within NCLB (2002) legislation.

Child behavioral engagement. Behavioral engagement is a construct defined by observable classroom behaviors that demonstrate the degree to which a child engages in his or her own learning experiences (Fantuzzo et al., 2007; Schaefer & McDermott, 1999). In the literature, this construct has also been called "behavioral regulation" (McClelland et al., 2007) "approaches to learning" (e.g., Denton & West, 2002; Denton, West, & Walston, 2003; Rathbun et al., 2005), "attention-related behaviors" (e.g., Claessens, Duncan, & Engel, 2009) and "learning-related social skills" (e.g., McClelland et al., 2000; McClelland et al., 2006), and it has been operationalized with slight variations, although the construct consistently includes a child's participation and cooperation in the classroom, self-control, attention to and persistence on tasks, and eagerness to learn (Alexander et al., 1993; Entwisle & Alexander, 1993; Fredricks,

Blumenfeld, & Paris, 2004; Greenwood, Horton, & Utley, 2002; Ponitz et al., 2009). For the purposes of clarity, the construct will be referred to as behavioral engagement in this research. Snow et al. (1998) noted the challenges relative to behavioral engagement for children within the kindergarten classroom environment:

> A child can no longer demand the attention or assistance of the attendant adult at will; each must learn how to solicit individual attention and to wait patiently while the teacher is attending to others. To a greater or lesser extent depending on the classroom, every kindergartner must learn to sit quietly, to listen considerately to both the teacher and other students, to communicate cooperatively, to restrain behavior to within acceptable limits, to accomplish tasks both independently and with others, to share resources, to treat others respectfully, and to try to learn and do what she or he is asked to learn and do (p. 179).

Poor behavioral engagement has been reported to occur disproportionately in children from at-risk backgrounds. For example, in a study examining differential outcomes between two groups of economically at-risk and not-at-risk 5- to 8-year-olds, the children did not differ in learning-related confidence or eagerness to learn, but children from economically at-risk families demonstrated poorer behavioral engagement skills, such as motivation and self-regulation (Howse, Lange, Farran, & Boyles, 2003). Yet, behavioral engagement can also act as a protective factor against high school dropout (Alexander et al., 1993; Entwisle & Alexander, 1993; Fredricks et al., 2004), and some have suggested that interventions targeting behavioral engagement skills might assist children in reducing the effects of the achievement gap (Fredricks et al., 2004). It is true that a child's behavioral engagement in academic tasks has been linked to later school achievement (Agostin & Bain, 1997; Alexander et al., 1993; Denton et al., 2003; Fredricks et al., 2004; McWayne, Fantuzzo, & McDermott, 2004) demonstrating predictive value (e.g., Benner, Beaudoin, Kinder, & Mooney, 2005; Malecki & Elliott, 2002) over and above the contribution of academic skills alone, for both at-risk and not-at-risk groups (Howse et al., 2003),

and enhancing the predictive ability of cognitive skills above and beyond that which cognitive ability and early achievement predict alone (Schaefer & McDermott, 1999; Yen et al., 2004). Yet, more research is needed to answer questions about the contribution of behavioral engagement skills in predicting later elementary school reading outcomes for struggling readers.

While the literature shows a clear influence of behavioral engagement on academic outcomes, the impact of this construct on reading-specific achievement is only beginning to be understood. Current literature shows positive within-year results for behavioral engagement skills predicting preschool emergent literacy and vocabulary skills (McClelland et al., 2007), predicting kindergarten reading skills (Ponitz et al., 2009), improving the predictive ability of kindergarten phonological awareness skills (Torgeson et al., 1999), and discriminating the growth of first grade reading scores above and beyond the predictive ability of first grade phonological awareness skills in a sample of children with poor reading development (Spira et al., 2005). However, while positive within-year results for kindergarten and first grade suggest that behavioral engagement skills may contribute to future reading scores, it is unclear whether these effects would continue throughout elementary school.

The few longitudinal studies examining the contribution of behavioral engagement skills on reading provide evidence for a relationship between behavioral engagement and later reading achievement from kindergarten through second grade (McClelland et al., 2000), fourth grade (Alexander et al., 1993), fifth grade (Claessens et al., 2009), and sixth grade (McClelland et al., 2006). Yet, within those studies, questions remain. First, although McClelland et al. (2000) reported that kindergarten behavioral engagement skills predicted unique variance in achievement at the end of second grade, and McClelland et al. (2006) reported that kindergarten

behavioral engagement skills were associated with gains in reading achievement from kindergarten through the second grade, McClelland et al. (2006) also reported that kindergarten behavioral engagement skills were not predictive of gains in reading skills beyond the second grade.

It is unclear from these results whether behavioral engagement skills are a less robust predictor of achievement as children progress through school, or whether kindergarten behavioral engagement skills lose significance as children grow and modify their behavioral repertoire. Another possible confounding factor in McClelland et al.'s (2006) study is their use of different measures after the second grade. Their results seem to defy expectations that as children progress through school and increase in independent learning requirements, their engagement in academic materials will become a more important indicator of success. No study, to date, has examined the influence of behavioral engagement skills, rated after kindergarten (i.e., first grade or third grade), on third and fifth grade reading outcomes to determine whether behavioral engagement, rated in first or third grade might better predict reading outcome scores at the end of elementary school.

Additionally, although current literature demonstrates a link between behavioral engagement and reading scores, studies examining this relationship tend to include an overall sample of children with diverse learning levels (Claessens et al., 2009; DiPerna, Lei, & Reid, 2007; McClelland et al., 2000; McClelland et al., 2006; McClelland et al., 2007), producing results that are appropriately generalized to the population of all elementary school students. Few have examined these questions with regard to samples of children who are likely to be targeted for academic reading skills interventions (Downer et al., 2007; Howse et al., 2003).

This distinction is important because children who are targeted for reading skills interventions in public schools tend to be associated with different demographic characteristics than is found in the general population of elementary school students, such as an increased likelihood of a being from a low SES background, coming from single parent households, and being from minority backgrounds (Rathbun et al., 2005). Consequently, the literature is in need of studies examining the predictive influence of behavioral engagement on elementary school reading outcomes in samples of children who might be selected for academic reading intervention so that empirical results are representative of the population of children who are targeted for intervention.

Additionally, in the behavioral engagement literature, to date, while children's early reading skills and kindergarten behavioral engagement skills have shown promise in predicting within-year and longitudinal outcomes, the potential confounding influence of children's participation in Reading First intervention has not been examined. This significant distinction may help to discriminate which specific characteristics (e.g., strong behavioral engagement skills) may assist children in benefiting more from school-based intervention. Without knowledge of a child's intervention status, variability in outcomes may be inaccurately attributed to a child's characteristics rather than the presence of effective school-based intervention. It remains unclear whether the influence of behavioral engagement skills could impact reading outcome scores above and beyond the combined influence of early reading scores and participation in Reading First programming.

The ECLS-K data provide a unique opportunity to examine the influence of behavioral engagement on later reading outcomes, using longitudinal data and consistent measures. This

research study expands on the current literature, further exploring these relationships by looking at the longitudinal predictive relationship of kindergarten, first, and third grade behavioral engagement ratings on reading outcome scores in third and fifth grades, using a nationally representative sample of students who were low performing in reading at the end of kindergarten.

Teacher influences. When examining the literature on early reading development, it is clear that multiple domains must be targeted to better assist children in the improvement of their academic reading skills. The usefulness of examining proximal processes surrounding teachers and instruction cannot be understated. A teacher's ability to modify instructional techniques in response to student cues is a powerful tool once referred to by Vygotsky (1978) as "scaffolding", and posited by Vygotsky to best occur within the Zone of Proximal Development (ZPD) - the distance between a child's capability to learn independently at their current developmental level and the child's skill development with assistance from an adult or more skilled peer. According to Vygotsky, children achieve higher developmental gains when proper scaffolding is included in their instruction.

However, NCLB (2002) legislation addressing reading intervention in elementary schools largely avoids discussion of instructional techniques. Rather, a main portion of the Act addresses teacher qualifications, or HQT characteristics. While HQT can be easily measured, their effectiveness in producing positive achievement gains for struggling readers is less clear. This research further explores the differential impact of teacher instructional style and HQT characteristics on elementary school reading outcomes in order to better inform policy efforts aimed at reducing the achievement gap.

Background characteristics. As part of a multivariate approach to address the goal of reducing and ultimately eliminating the achievement gap, NCLB contains regulations defining characteristics of a HQT (20 U.S.C. § 7801 [2002]), and regulations requiring that all teachers providing Reading First services meet the HQT requirement (20 U.S.C. § 6362 [2002]). Generally speaking, these regulations mandate that teachers are certified by the state and possess knowledge in the subjects they teach. While certification status has been linked to student achievement (Chatterji, 2006; Kukla-Acevedo, 2009; Goldhaber & Brewer, 1997; Goldhaber & Brewer, 2000; Hawk et al., 1985; Heck, 2007; Nye et al., 2004; Wayne & Youngs, 2003), there have been no longitudinal studies examining the effect of certification on the reading achievement gains of struggling readers. Additionally, some previous work used school-level analyses, such as Chatterji's (2006) study showing strong school-level achievement effects of certification rates on first-grade reading achievement, and Heck's (2007) study showing a positive relationship between teacher quality and reading achievement in elementary school, particularly for children from risk groups. While these results are informative, neither study can link their findings to individual students. Other studies found a link between teacher certification and within-year math scores of students in fifth grade (Kukla-Acevedo, 2009), tenth grade (Goldhaber & Brewer, 1997), and twelfth grade (Goldhaber & Brewer, 1997). While this, too, provides valuable information about teacher certification effects on achievement, they do not address whether similar effects exist for the reading achievement of struggling readers. No study, to date, has examined longitudinal data of HQT status on the reading outcomes of elementary-aged struggling readers.

Others findings have suggested that there is a lack of evidence to support policy linking HQT characteristics with student achievement (Early et al., 2006; Early et al., 2007; Guarino, Hamilton, Lockwood, Rathbun, & Germino-Hausken, 2006; Muñoz & Chang, 2007; Rowan, Correnti, & Miller, 2002), using results based on samples of preschool children (Early et al., 2006; Early et al., 2007) and high school children (Guarino et al., 2006; Muñoz & Chang, 2007). Of the few studies using elementary school samples, results were somewhat inconsistent. Rowan et al. reported that neither special certification to teach reading, nor a bachelor's or master's degree in English, had a significant effect on growth in first-grade children's reading achievement; however, their results were based on a relatively small sample of teachers.

Another study reported no significant effects of certification status, but positive effects for type of degree earned (i.e., elementary education; Croninger, Rice, Rathbun, & Nisho, 2003). However, neither study used a sample struggling readers. Further, while both studies reported no significant link between teacher certification and child reading achievement, their results differed on the question of whether teacher degree makes a significant contribution to child reading achievement. Consequently, while the HQT regulations remain a focus for school systems aiming to comply with NCLB regulations, a consensus has yet to emerge in the literature regarding the effectiveness of these characteristics on the reading achievement of at-risk elementary school students.

To date, no study has examined the differential effect of HQT characteristics on the reading outcomes of struggling readers in elementary school who are in need of better developed academic reading skills upon school entry - the population of students targeted with Reading First programming. Further, the field would benefit from studies examining these questions using

longitudinal data and nationally representative samples of children who reflect the group of students targeted with Reading First legislation (NCLB, 2002). This research further extends the literature by examining the effect of teacher certification, and longitudinally, from kindergarten through fifth grade with a nationally representative sample of kindergarteners who are struggling with reading achievement. Additionally, the type of degree earned will be examined to clarify the issue of whether the influence of degree status extends to this specific population of children.

Instruction. In addition to teacher characteristics, regulations in NCLB (2002) focused on increasing the instructional time dedicated to the five essential components of reading. The literature in this area is inconsistent with regard to the effectiveness of increased reading instruction time on student reading achievement outcomes. While increasing time on instruction has been associated with improved academic performance (Connor, Son, Hindman, & Morrison, 2005), others found that increased time was not enough to increase reading comprehension skills (Block et al., 2009). Some posit that positive outcomes of increased time on reading instruction can only occur when effective, empirically-validated instructional techniques, are used to improve student outcomes (Block et al., 2009).

One area of the literature that has shown positive results is examining the contribution of teachers instructional choices on student achievement (Block et al., 2009; Mashburn et al., 2008; Parlardy & Rumberger, 2008; Stronge, Ward, Tucker, & Hindman, 2007). Recent research suggests that teachers instructional choices can substantially influence student outcomes, perhaps beyond that of HQT characteristics (Mashburn et al., 2008), or that of increased instructional time (Block et al., 2009; Parlardy & Rumberger, 2008).

In this regard, Mashburn et al. reported that teacher's use of instructional techniques such as integration of curricula and discussion demonstrated a significant contribution to achievement outcomes for pre-kindergarteners; however, these results were based on a sample representing the overall population rather than children who are struggling with reading skills. Further, Mashburn et al.'s results examined overall achievement and did not specifically address reading achievement. Extending their results to older elementary-aged children will be beneficial in clarifying whether instructional techniques such as discussion and integrating curricula demonstrate positive achievement gains for older elementary aged children.

Block et al. (2009) attempted to further explore some of these questions by examining children's reading comprehension outcomes in grades two through six and found that children who were given more instructional time without the addition of teacher instructional interactions - such as choice and discussion - produced the lowest gains in reading comprehension. However, the sample in Block et al.'s study represented the general population of elementary students, and thus cannot be generalized to the population of elementary children targeted through NCLB (2002) legislation.

As a result, questions arise as to whether the benefit of increased instructional time on student gains is linked specifically to instructional techniques and whether these benefits extend specifically to struggling readers. Current literature points to the use of instructional techniques that encourage student interaction and engagement with the academic materials, such as choice, integrating curricula, and discussion to produce positive gains in student achievement (Block et al., 2009; Connor et al., 2005; Hamre & Pianta, 2005; Houtveen & van de Grift, 2007; Mashburn et al., 2008; Pressley et al., 2001; Rowan et al., 2002; Stronge et al., 2007; Wenglinsky, 2002).

When considering the effect of instructional practices on the general population of students, teachers' instructional choices have been linked with improved achievement in elementary students (Houtveen & van de Grift, 2007; Stronge et al., 2007), and middle school students (Wenglinsky, 2002). Further, Xue and Meisels (2004) reported that higher reading outcomes in kindergarten were linked to teachers' instructional techniques (e.g., discussion, cross curricular connections, and choice) in a general population sample.

The literature is less clear about whether the impact of teachers instructional choices remains as strong in samples of children with less well-developed reading skills. One author reported that teachers with high proportions of kindergarten and first graders from low-income backgrounds used basic skill instruction techniques (e.g., rote memorization, reading without discussion, isolated phonics tasks without meaningful text, etc.) more frequently than engaging instructional techniques (e.g., discussion, phonics embedded in meaningful text; emphasis on student participation; Stipek, 2004); however, their results did not examine whether children demonstrated differential achievement outcomes.

Interestingly, Torgeson (2002) reported that children from at-risk backgrounds, such as low SES, benefited more from basic-skills instruction over higher-order, pedagogical techniques. Additionally, Xue and Meisels (2004) reported that at-risk kindergarten-aged children with poor achievement were found to benefit less from instructional techniques (e.g., discussion about reading, cross-curricular connections, dictating and retelling stories) than children who were not at-risk, suggesting that children who are at-risk may benefit more from basic-skills instructional practices than would children who are not at-risk.

However, Hamre and Pianta (2005) challenged Xue and Meisels' (2004) results and Torgeson's (2002) results, reporting improved overall achievement gains for children who were at-risk and learned from teachers who reported using instructional techniques such as evaluative feedback, discussion, and encouragement of student responsibility, thus suggesting that children who are at-risk may benefit from engaging instructional practices. It is unclear from these results whether children who are considered at-risk for poor achievement differentially benefit from exposure to engaging instructional techniques. The literature is certainly in need of clarification on this question.

Another body of literature exists that specifically examines the literacy practices of children who have scored poorly on standardized reading tests, and are therefore considered to be at-risk for poor achievement. Findings indicate that when these children are out of school and provided with choice and appealing activities that include reading, many children will choose to engage in these alternate literacy-type activities (e.g., texting, using the Internet) and will self-identify as "readers" (Alvermann et al., 2002; Alvermann et al., 2007). The authors noted these results as somewhat unexpected, given the students self-identification as being uninterested in reading.

However, Alvermann et al's (2007) results may better explain inconsistencies in the instructional practices literature (Connor et al., 2005; Hamre and Pianta, 2005; Torgeson, 2002; Xue and Meisels, 2004) and also provide further insight into what types of practices could best assist children in improving their academic literacy skills in public schools. For example, given Alvermann et al.'s (2007) findings, it would seem that instructional practices that incorporate student choice of materials and involve teaching students to engage with others regarding

reading materials may better parallel the out-of-school literacy practices of this sample of students and lead to more positive in-school reading outcomes. These questions surrounding the impact of choice, discussion, and cross-curricular instructional practices are important to examine within the framework of struggling readers.

Although results from the literature demonstrate inconsistent results in determining whether children from at-risk backgrounds benefit from their teachers' differential use of these instructional practices, the use of overall achievement outcomes, rather than reading achievement, may contribute to the lack of consistent results. Differences in effective instructional practices across subject areas can be vast - what might work well in mathematics could be less effective in reading instruction. Focusing results on a the reading achievement of struggling readers could better clarify the impact of these instructional practices on the achievement of children who are at-risk. Additionally, the current literature examining teachers' instructional practices is absent of clear longitudinal data across elementary school. Likewise, only a few studies examining instructional effects on achievement have focused on the population of students targeted through Reading First legislation, and their results have been inconsistent (Connor et al., 2005; Hamre & Pianta, 2005; Torgeson, 2002; Xue & Meisels, 2004).

There is a need to expand the literature using longitudinal data of kindergarten and elementary-aged students who have less well-developed reading achievement in school, paying particular attention to the effects of early skills and early instruction on student achievement in later elementary school. The ECLS-K data provide a unique opportunity to use a nationally representative sample of kindergarten students and teachers who are followed with consistent measures throughout elementary school. With these data, it was be possible to examine the

unique effects of teacher instructional style on young children who are at risk for poor academic reading achievement, and determine whether differential effects exist between teacher qualifications and instructional style on struggling readers achievement in third and fifth grade.

Relation Between Behavioral Engagement and Teacher Influence

While teacher's instructional practices that support engagement can work to bi-directionally support children's ultimate reading achievement (Connor et al., 2005; Ponitz et al., 2009), even teachers using instructional approaches strongly supported in the literature can only be successful if met by children who can focus and engage with the tasks at hand (Downer et al., 2007; Hughes & Kwok, 2007). Ponitz et al. (2009) described the interaction between teaching qualities and a child's behavioral engagement as follows:

> [teachers] actively monitor and respond to child engagement. Compared to learning gains, behavioral engagement provides a relatively proximal indicator of whether instructional efforts are succeeding…. If students are not engaged, teachers can then adjust their actions to promote these desired behaviors. (p. 117).

Brophy and Good (1986) discussed methods teachers can use to elicit student engagement, making a connection between teachers actions and student responses. More recently, Ponitz et al. (2009) examined child behavioral engagement as a mediating factor between classroom quality and child reading achievement. For the purposes of their study, classroom quality was measured explicitly through observation of teacher behavior and covered emotional support, classroom organization, and instructional support. Their findings revealed that while classroom quality, as defined by teacher behavior, had an indirect effect on student achievement; this effect was mediated by child behavioral engagement. These results suggest an interactive effect of child behavioral engagement and teacher behavior on child achievement

outcomes. However, Ponitz and colleagues' results were within year and generalizable only to an overall sample of kindergarten children. Further, their measures of teacher behavior covered multiple domains of behavior and results on subdomains (e.g., teacher instructional support) were not reported.

The field is in need of further clarification about the interaction between child behavioral engagement and specific teacher variables (i.e., instructional style and HQT status). If child behavioral engagement is found to contribute, significantly, to student reading outcomes in third and fifth grades, then determining what teacher variables, if any, best produces high engagement from elementary students will be helpful to inform teaching practices when working with struggling readers. This research provides a preliminary exploration of this question by determining whether behavioral engagement scores at points across elementary school (kindergarten, first, third, and fifth grade) had any combined effect with teacher instructional style during the same school year, and whether these skills have any effect on student reading outcomes. These results should be used to inform future research in this area.

CHAPTER III: RESEARCH STUDY

Given the importance of NCLB (2002) legislation in combating educational inequities, particularly those in reading development, identifying school variables, above and beyond a child's background, that best assist school-aged children in achieving reading proficiency is paramount. Applying bio-ecological theory (Bronfenbrenner & Morris, 1998) to child reading development, it follows that reading develops over time through a series of interactions between a child's personal and background characteristics and resources or materials available to the child (including the child's parents and teachers). Deficiency or poor development in any area can have a compounding effect on the child's development (Rathbun et al., 2005; Zill & West, 2001), thus highlighting the importance of strong resources and effective teaching strategies, particularly for children who enter formal schooling without adequate pre-reading skills. Furthermore, with strong supports and resources, focusing on developing a positive repertoire of child behavioral characteristics in school, such as engagement in classroom activities, may serve as a protective factor in the challenging pursuit of reading proficiency.

NCLB (2002) legislation targeting the achievement of reading proficiency highlights a multifaceted intervention to better develop child reading skills, but it does so through a narrow lens: child characteristics are primarily limited to five reading-specific skills (phonemic awareness, phonics, vocabulary, fluency, and comprehension), and teacher quality is primarily limited to certification and subject-matter knowledge. The Act is largely absent the inclusion of

interactions that occur as a culmination of background variables such as behavior, engagement in school activities, and teacher-child interactions (i.e., process variables; Bronfenbrenner & Morris, 1998). In other words, a struggling reader, who is also unable to follow directions or focus and persist on classroom assignments may have a far more difficult time developing age-appropriate reading proficiency than a struggling reader who has strong classroom behavioral engagement skills (Alexander et al., 1993; McClelland et al., 2000; McClelland et al., 2006; Ponitz et al., 2009; Spira et al., 2005).

Current literature demonstrates a link between the child's classroom behavioral engagement and academic achievement. However, the literature largely depends on results using samples of children with a full range of reading skills levels, rather than the those students who are already demonstrating poor achievement. Additionally, the current literature tends to focus on overall achievement outcome scores, rather than reading-specific outcomes. Further, conclusions about the utility of behavioral engagement as a predictor of later achievement are based on kindergarten behavioral ratings. No study has yet examined the effect of behavioral characteristics throughout elementary school, on reading achievement scores. A connection between child behavioral engagement and reading proficiency would suggest that reading interventions should be developed within the framework of encouraging and maintaining adaptive behavioral engagement skills in order to improve a child's ability to benefit from instruction.

Likewise, a struggling child who has been given increased instructional time with a certified teacher may not benefit from certain types of instruction, thus demonstrating inadequate gains in reading (Block et al., 2009; Hamre & Pianta, 2005; Houtveen & van de Grift, 2007;

Mashburn et al., 2008; Parlardy & Rumberger, 2008; Pressley et al., 2001; Rowan et al., 2002; Wenglinsky, 2002; Xue & Meisels, 2004). Indeed, a body of literature highlighting the influence of a teacher's instructional strategies on childhood academic achievement, suggests that instructional influences can be stronger than that of teacher certification alone (Mashburn et al., 2008). Yet, no studies have examined the effects of teacher certification and teacher instructional strategies with an elementary school population struggling readers. Process variables such as behavioral engagement and teacher instructional style, are a framework that can be used to improve the quality and outcomes of reading interventions funded through NCLB.

Research Hypotheses

Based on findings from previous research regarding the application of child behavioral engagement and teacher instructional practices to child reading outcomes, the following hypotheses were proposed for this research:

1. Kindergarten, first, third, and fifth grade child behavioral engagement skills will significantly add to the prediction of child literacy outcome in third and fifth grade beyond the individual contribution of children's kindergarten literacy skills and the combined contribution of their kindergarten literacy skills and participation in Reading First interventions.

2. Differences in teachers' instructional strategies will significantly add to the prediction of children's third and fifth grade literacy outcome beyond the contribution of teachers' HQT status alone, and beyond the combined contribution of teachers' HQT status and children's participation in additional instructional time.

3. Child behavioral engagement and teacher instructional style will demonstrate within-year interactions and individual effects on reading outcomes at points across elementary school (kindergarten, first, third, and fifth grades).

Methods

Participants. Data utilized for this research came from the ECLS-K fifth grade public-use longitudinal data file (NCES, 2006). The ECLS-K is a federally funded longitudinal study that focuses on children's early school experiences, following children from kindergarten through eighth grade. In the base-year, the ECLS-K sample was a nationally representative sample of approximately 21,260 children who attended kindergarten from 1998 to 1999 and were selected through a complex, clustered multi-stage probability design. For a complete description of the base year selection procedures for the ECLS-K, please see Tourangeau et al., (2004). Attrition from the kindergarten year to fifth grade occurred largely due to a child's family moving to an non-sampled location, hard refusals[1] by a child's parents, or a lack of any data in both first and third grade. Figure 1 displays a general overview of the following methods, which detail how the analytic sample was developed from the fifth grade public-use longitudinal dataset.

The fifth-grade longitudinal dataset[2] is composed of 10,673 eligible participants (51.5% male, 48.5% female) with an average age of 6.23 years in the spring of kindergarten year. Participants were from diverse backgrounds (57.7% White, 18.8% Hispanic, 16.0% Black, 2.9% Asian, 1.7% Native American, 2.3% Multiracial, 0.6% Pacific Islander)[3]. Eighty percent of eligible participants attended public school across all for years, and 96% of eligible participants were first-time kindergartners. Of eligible participants, 74.4% lived in a home with two parents,

[1] Hard refusals is defined in the ECLS-K literature as "children whose parents emphatically refused to cooperate" (Tourangeau, Nord, Pollack, & Atkins-Burnett, 2006 p. 4-2)

[2] These data exclude children from the Round 3 freshening sample (n = 5118). Please see Tourangeau et al., (2002) for more information about the excluded sample.

[3] These are weighted estimates. Please see Section 4.7 in Tourangeau et al. (2006) for a complete description of weights in the ECLS-K.

while 24.1% reporting living in a single parent home, and 1.5% reported living with no biological or adoptive parent in the home. Further 12.6% of participants reported that a non-English language was spoken as the primary language in their homes at the time of kindergarten data collection, one percent reported difficulty hearing, and 3.6% reported difficulty with vision. Twenty-two percent of eligible participants lived below the poverty level in kindergarten, and 34% reported participation in the free school lunch program[4]. During the kindergarten year, 4.5% of eligible participants were involved in special education services, and 10.5% participated in Reading First reading services.

Sampling procedures. Prior to developing the analytic sample, multiple imputation (MI) procedures were completed to recover missing data, as discussed in the *Missing Data* section later in this chapter. An effort was made in this research to accurately capture the literacy skill development of children who attend public elementary schools and began formal schooling with poorly developed academic pre-literacy skills. As a result, a number of exclusions, outlined in Figure 2, were made to the sample of eligible children. First, less than 0.3% of participants (n = 22) were excluded from analysis because of teacher reported participation in an ungraded classroom for the majority of the school day, as this type of classroom experience often provides an environment and curricula that differs from a typical public elementary school experience.[5] The sample was then limited to those who were first-time kindergarteners during Spring, 1999 (kindergarten) data collection (n = 10279) and who attended public school during all four rounds

[4] Children are eligible for this program if their household income is less than 130% (free lunch) or 185% (reduced lunch) of the federal poverty guidelines. For more information, please see the Program Fact Sheet (Food and Nutrition Service, 2013)

[5] A description of how these participants were excluded from analysis can be found in the *Results* section of this document, under the sub-heading, *Child Variables.*

of data collection (n = 8047). Children with exposure to a kindergarten classroom prior to the first round of data collection were not appropriate for the analyses, as initial estimates of their knowledge would include previous kindergarten-level experiences, potentially skewing initial data estimates. Further, limiting the sample to children who attended public school during each round of data collection ensured that the analyses captured an accurate picture of literacy achievement within a public school environment.

Of the remaining sample, additional participants were excluded due to parent report of a primary language other than English (n = 1424), and/or a non-correctable hearing (n = 60) or vision (n = 232) problem diagnosed before the questionnaires were completed in the fall of kindergarten data collection. As previously noted, participants who skipped grades or were held back after the kindergarten year were retained in the sample. Following these exclusions, 6331 participants remained eligible for analysis. Finally, the sample was limited to participants who scored at or below the 30th percentile (n = 34.062) on IRT-based theta estimates of overall academic reading skills during kindergarten data collection. Table 1 displays the data used in the calculation of the IRT cutoff score, calculated by taking the mean of the 30th percentile cutoff score for each of the 20 imputed datasets[6]. Following this final exclusion, the sample contained 1880 participants. Hereafter, rounds of data collection will be referred to by the grade in which the majority of children were enrolled. For example, the spring, 1999 data collection will be referred to as "kindergarten data collection".

[6] Please refer to the *Missing Data* section later in this chapter for a description of MI procedures that resulted in these data.

Sampled participants. Sampled participants (n = 1880; 57.4% male, 42.6% female) were first-time kindergartners in public schools throughout the United States (US), representing the northeast (17.5%), south (43.4%), midwest (24.3%), and west (14.8%), as well as US cities (35.2%), suburbs (38.8%), and rural areas (26.0%). Male participants were of similar age (M= 6.15 years) to the female participants (M = 6.14 years,). Participants were from diverse backgrounds (50.0% White, 30.2% Black, 12.6% Hispanic, 1.1% Asian, 0.7% Pacific Islander, 3.4% Native American, and 2.1% Multiracial), and, as expected, many lived in circumstances of increased risk for poor academic performance (Coie et al., 1993). For example, at the time of kindergarten data collection, 35.2% of sampled participants were living below the poverty level, 2.7% were living in a home with no biological or adoptive parent, and 37.6% were in a one parent household. Further, at the time of kindergarten data collection, 47.9% of sampled participants were involved with the school free lunch program, 9.2% were receiving special education services, and 17.8% were receiving Reading First services.

Table 2 describes how sampled participants differed from non-sampled participants across many areas of risk including age, race, SES, Reading First participation, number of parents in the home, special education participation, and retention in school. Consistent with previous literature, these data demonstrate how children who begin school with poor pre-literacy skills differ in background characteristics from what might be expected in the general population of students (Downer et al., 2007; Howse et al., 2003; Rathbun et al., 2005; Vitaro et al., 2005).

When examining the academic pre-literacy and literacy skills of participants, significant differences were found between the kindergarten skill development of sampled and excluded participants in letter recognition, beginning sounds, and ending sounds. Specifically, Table 3

shows that kindergarten children who had not yet mastered letter recognition, and those who had mastered letter recognition but not yet mastered beginning sounds, were significantly more likely to be included in the analytic sample. Kindergarten children who had mastered beginning sounds had an equal chance of being included in or excluded from the sample. Finally, those who had mastered ending sounds by kindergarten were significantly more likely to be excluded from the sample, scoring above the 30th percentile on kindergarten Reading IRT scores. At kindergarten entry, only 1.6% of sampled participants had mastered ending sounds compared with 43.8% of excluded participants. Similarly, 52.5% of sampled participants had mastered letter recognition but not yet achieved mastery of beginning or ending sounds compared with 11.1% of excluded participants at the same level of skill mastery. In other words, participants were much more likely to be sampled in kindergarten if he or she had not achieved mastery of skills beyond letter recognition or if he or she had not yet mastered ending sounds.

Teachers of sampled participants. While it is important to understand the characteristics of participants' teachers, the data reported in this paper describes that of each participant's teacher, with the understanding that a large percentage (ranging from 43% for fifth grade teachers to as many as 63% for kindergarten teachers) of teachers completed questionnaires for more than one sampled participant. Although, most teachers from kindergarten through fifth grade completed three or fewer questionnaires (ranging from 76% of kindergarten teachers to 90% of third grade teachers), it is imperative to only interpret these data at the child level. Table 4 outlines the frequency and percentage of questionnaires completed by teachers. Data describing participants' teachers cannot be viewed as data describing individuals, but rather as a description of individual participants' teachers. For example, a teacher with three sampled

participants in the classroom would have completed one questionnaire for each participant and consequently the teacher data would describe that teacher three times. Descriptive statistics for sampled participants' teachers, across each year of data collection, are presented in Table 5. Notably, sampled participants' teachers were, as a group, relatively similar across years of data collection. The comparatively low variability in the percentages of participants' teachers certification status will be discussed in the Results section.

Procedures.

Approval by the human subjects committee. This study complied with ethical issues and standards set forth by the American Psychological Association and the University of North Carolina at Chapel Hill. Approval was obtained from the Institutional Review Board of the University of North Carolina at Chapel Hill in April, 2010 (ID: 10-0639).

Materials. Secondary data derived from direct child assessments, teacher interviews and parent interviews across four rounds of data collection from spring, 1999 through spring, 2004 (i.e., spring of kindergarten, first, third, and fifth grade) was used to examine the research questions. The ECLS-K was sponsored by the U.S. Department of Education, NCES, Institute of Education Sciences and conducted in collaboration with other federal agencies, including Office of Special Education Program (OSEP) and the Administration on Children, Youth, and Families (ACYP). For a complete list of the sponsoring and collaborating agencies, please visit the ECLS-K website (http://nces.ed.gov). Each of the measures used in this study are available to the public via the ECLS-K website, with the exception of the Social Rating Scale (SRS), due to copyright reproduction restrictions.

Child characteristics and demographic information. Participants' background characteristics and demographic information was obtained from parent/caregiver report during computer-assisted interviews conducted at the time of kindergarten data collection. Descriptive variables including participant family status, race, gender, SES, child health status (e.g., vision and hearing diagnoses), and primary language spoken in the home were obtained using these methods.

Child literacy skills. Direct literacy assessments, designed to measure cognitive knowledge and skills representative of typical elementary curricula, were individually completed with all eligible children during each round of data collection, using both hard-copy instruments and computer-assisted interviews (Pollack, Atkins-Burnett, Najarian, & Rock, 2005). Items were designed to "sample typical and important elements of the curriculum with particular emphasis on content and process areas that are critical to growth and can be expected to reflect growth on the same scale over time" (West et al., 2000, p. 11). The literacy assessment during the kindergarten and first grade rounds specifically focused on basic early literacy skills, such as letter recognition and vocabulary. During third and fifth grade rounds, more advanced skills were added to the assessment, including phonemic awareness, passage comprehension, and inference. For a detailed description of administration procedures for the instrument, please see Tourangeau, et al (2006).

Reliability for estimates (IRT-based theta[7]) of overall reading ability was over .90 during each round of data collection (Pollack et al., 2005). Construct validity using the Woodcock-McGrew-Werder Mini Battery of Achievement (MBA; Woodcock, McGrew, & Werder, 1994)

[7] Please refer to Tourangeau et al. (2006) for a complete description of ECLS-K reading IRT scores.

supported the validity of the literacy test items with moderate to high correlations of .83 for the third grade measure and .73 on the fifth grade measure (Pollack et al., 2005). The decrease in correlation as participants progressed through school was attributed to content differences in the measures such as a greater emphasis on comprehension and higher order skills as participants grew older (Pollack et al., 2005, p. 2-22). To reduce the overall length of the assessment to an average time of 75 minutes, the test was conducted in two stages: (1) a routing test of 26 items to determine appropriate level of difficulty; and (2) a set of actual test items (between 16 and 40, depending on difficulty level) centered around passages the children were asked to read prior to answering the questions.

ECLS-K participants were assigned both normative and non-normative scores from each round of literacy assessments. Overall IRT-based scores collected in the spring of participants' kindergarten year were used to develop the initial sample, for this research, while subsequent IRT-based scores (first, third, and fifth grade) were used to estimate sampled participants' reading skills later in elementary school. Although comparisons of IRT scale scores between rounds can have different interpretations depending on the actual test items completed, the analyses were considered meaningful due to the similar initial status (i.e., scores at or below the 30th percentile) of participants' kindergarten IRT scores (Tourangeau et al., 2006).

Teacher ratings. During each round of data collection, teachers completed self-administered questionnaires in which they disclosed their personal demographic information, professional background, and teaching practices. While best practice suggests incorporating data from direct observations of teacher's instructional practices, these types of data were unavailable in the ECLS-K due to time and monetary restrictions (Tourangeau et al., 2006). In lieu of direct

observation, the literature supports the use of teacher self-report. Elementary school teachers' self-reports of their instructional practices have been found to be remarkably consistent with third-party observations (Stipek & Byler, 2004). These findings suggest that the data available in the ECLS-K can be considered reliable and valid reports of teachers' instructional practices.

Teachers of eligible children also completed a questionnaire for each participating child in their classrooms, covering content related to academics (reading, mathematics, or science) and classroom behavior. Data from independent variables such as participants' involvement in Reading First (rated dichotomously, yes or no) and behavioral engagement skills (rated on a Likert-type scale, ranging from 1 to 4) were obtained from this measure.

The final portion of the teacher questionnaire was adopted, with permission, from a previously published instrument, *Elementary Scale A "How Often?"* (Gresham & Elliott, 1990b), and was designed for teachers to rate the frequency of specific child social skills and behaviors. Referred to as the Teacher Social Rating Scale (T-SRS) or more generally, the Social Skill Rating System (SSRS; Gresham & Elliott, 1990a), it is considered a reliable and valid measure of children's social development (Tourangeau et al., 2006), and is regarded as a comprehensive rating system of children's social development because of the use of multiple sources and direct links to intervention (Demaray et al., 1995), and. One significant advantage to using the SSRS is how it differentiates school-related social skills from other social skills (Vaughn & Haager, 1994). This distinction highlights the unique opportunity presented by the use of ECLS-K data to examine the effect of children's school-related social behaviors, such as their classroom behavioral engagement on their achievement outcomes.

Behavioral engagement ratings on the T-SRS consisted of six to seven questions intended to measure children's attention, persistence on tasks, eagerness to learn, independence in learning tasks, flexibility, organization, and classroom rule following (during third and fifth grade data collection). Table 6 presents summary statistics for each of the variables used to create kindergarten, first, third, and fifth grade behavioral engagement scores. For each item, teachers rated the students on a Likert-type frequency scale ranging from Never (1) to Very Often (4). The mean rating of the items was used to compute participants' overall scores for the ECLS-K dataset. Mean scores were only computed by the ECLS-K if the participants' teacher provided a rating for at least two thirds of the items. According to data reported in the ECLS-K manuals, the behavioral engagement sub-scale — called Approaches to Learning in the ECLS-K — was a reliable (split-half reliability coefficients ranging from .89 in kindergarten to .91 in fifth grade) and valid measure of children's social skills (Pollack et al., 2005).

Research Design.

Data were assessed through the ECLS-K fifth grade public-use longitudinal data file (NCES, 2006), which employs a complex, clustered multi-stage probability design, using primary sampling units (PSUs), or "geographic areas consisting of counties or groups of counties. The second stage units were schools within sampled PSUs. The third and final stage units were students within schools." (Tourangeau et al., 2006, p. 4-1.). All participants from the kindergarten year were eligible for participation in subsequent years if there was a completed child assessment or a parent interview from the kindergarten year. For a detailed description of the selection procedures for the ECLS-K, please see Tourangeau et al., (2004). Additionally, oversampling was used to facilitate the study of some underrepresented populations, such as

people of Asian and Pacific Island descent as well as participants who moved and children whose home language was not English (Tourangeau et al., 2006).

Eligible participants for this research (n = 10673) were sampled from 89 PSUs and 538 schools within those PSUs. The number of schools within each PSU ranged from 2 to 75, although all but two of the PSUs contained 12 or fewer schools. Close to 35% (n = 31) of the PSUs contained only two schools, and overall, 87.6% (n = 78) of PSUs contained fewer than ten schools.

Design effects. Due to the large sample size and nationally representative nature of the data on variables such as race, gender, and socioeconomic status, the effects of random error and bias were minimized. Maturation was not a threat to internal validity because the research questions specifically addressed social and academic development over time.

Use of weights. Due to the complex nature of the design of the ECLS-K, the use of probability weights was necessary to adjust for the effects of non-response and for the differential probabilities of selection at each sampling stage (Tourangeau et al., 2004). The use of weights produced estimates that were representative of the population of children who entered kindergarten in 1998-1999. The ECLS-K weight, c*2_6fc0* (n =10673, M = 359.60, SD = 596.79, range 1.75 to 6360.58) was applied to the analyses because it was created to be applied to child and teacher data from the spring rounds of kindergarten, first, third, and fifth grade data collection. Detailed information about the use of weights in the ECLS-K and about the weight chosen for the analyses are described by Tourangeau et al. (2006).

Survey data. The ECLS-K weights and variance estimation variables were applied to the data using survey commands within STATA SE, version 11 (StataCorp., 2009b). Using the longitudinal weight *c2_6fc0*, PSU variable, *c26fcpsu*, and strata variable, *c26fcstr*, the data were set to correctly apply the weights in analyses using the survey commands, as well as to provide appropriate estimations of the sample variance.

Variance estimation. When using a complex sample design, one cannot assume simple random sampling in the calculation of variance estimates (Rao, 1996). With a nested survey design, an assumption of simple random sampling would lead to underestimated standard errors and could result in incorrectly assigning statistical significance to results. As a result, the Taylor Series estimation method was employed, as it was designed to estimate standard errors in complex sample designs. In a complex multi-stage design such as ECLS-K, the Taylor series method estimates variance from the variance among PSUs and combines stratum variance estimates to compute the overall variance estimate (Lee & Forthofer, 2006; Wolter, 2007). STATA SE, version 11.2 has the capabilities to perform all data management, multiple imputation, all planned analyses, and the Taylor series variance estimation using both the multiple imputation and survey data commands (StataCorp., 2009a). Therefore, it was determined that STATA SE would be used for all statistical procedures, rather than SPSS as proposed.

Data Preparation

Extensive data preparation was necessary to ensure all data were coded accurately for analysis. First, missing values were cross-referenced and recoded, as necessary. Categories within variables were examined and collapsed to clarify the data for analysis. Additional

variables necessary for analysis were created by combining variables together, as explained in the "variable creation" and "composite variables" sections. Finally, missing data were recovered using MI procedures discussed in the "missing data" section. Unless otherwise described below, missing values were cross referenced with ECLS-K questionnaires and re-coded as "system missing" (.), the only acceptable coding for inclusion in STATA MI analysis.

Missing data coding. As discussed in Tourangeau and colleagues' (2006) work, ECLS-K data contained a series of codes to categorize types of missing values: system missing (.), refused (-7), don't know (-8), not ascertained (-9), and not applicable (-1) . The coding, "system missing", was most common and typically indicated that an entire instrument or assessment was missing from the data set. All data coded as "system missing" were retained as "system missing"across all variables for inclusion in MI analysis. Responses "refused" and "don't know" indicated that respondents either did not want to answer specific questions or did not know the answer to a question; these responses were recoded to "system missing" across all variables. Values coded, "not ascertained", indicated that the respondent left a question blank that should otherwise have been answered, for example, unintentionally skipping a question. These responses were also recoded to "system missing" across all variables.

Finally, data coded, "not applicable", typically indicated nonresponse due to a planned skip pattern within the questionnaire; however, examination of the skip patterns revealed that these results were often meaningful for analysis. As a result, these data were cross-referenced with questionnaires and other individual variables to determine how they might best be recoded. Table 7 presents the frequencies of "not applicable" codes within the analytic dataset, along with an explanation for how the values were recoded for each variable.

Collapsing categories. Categories within participant variables such as race, grade level, school changes, family composition, and participation in Reading First programming, as well as variables describing participants' teachers' education, certification, teaching preparation, and instruction choices were collapsed to either reflect the level of information desired to answer the research questions, or to combine categories if the within-sample cell frequencies dropped too low (n = 50).

Participant variables[8]. The ECLS-K child race variable consisted of eight categories, two of which were reserved for people of Hispanic origin who were multiracial (5.8%) or who were Hispanic without another defined race (6.6%). The two Hispanic categories were collapsed into one category labeled "participants of Hispanic origins". Next, each of the racial categories were cross-referenced with each other to ensure that they were mutually exclusive categories. Participants who appeared in more than one racial category were recoded as "multiracial". A total of seven categories of race (White, Black, Hispanic, Asian, Pacific Islander, Native American, and Multiracial) were included in analyses.

All sampled participants, by design, were first time kindergartners at the time of the spring kindergarten data collection. However, over time, some of these participants were retained in grades or skipped grades, resulting in multiple grade categories for first, third, and fifth grade data collection. For clarity during analysis, these data were recoded to reflect the grade of the child (0) kindergarten (1) first grade (2) second grade ... through (8) eighth grade. Additionally, in the first grade variable, or when most (n = 10277, 96.3%) of eligible participants were in first grade, categories reflecting different types of kindergarten: half-day (n =47, < 0.01%), full day

[8]All data frequencies reported in this section are unweighted estimates of the original data (M = 0) with missing values because they explain facets of variables that were collapsed prior to MI procedures.

(n = 177, < 0.02%), and half/ full day unknown (n =96, < 0.01%), were collapsed into one kindergarten category (n = 320, 0.03%). Participants who were coded as learning in ungraded classrooms (n = 22) were recoded to "hard missing" (.a) because this was the most appropriate code for inclusion of these data in the MI dataset without recoding them as "system missing", which would have resulted in imputed values. As a result, these cases were excluded from the analytic sample (see Figure 2) prior to analysis.

Participant mobility between schools was recorded at the time of first, third, and fifth grade data collection with regards to whether a school change occurred since the previous data collection point. The original school change variables contained six categories reflecting whether the change in schools was between any combination of public and /or private school. However, because participants not in public school were sampled out of these analyses only two categories were retained for analysis: no school change, coded (0) and change between two public schools, coded (1).

ELCS-K data describing participants' family composition at the time of kindergarten data collection included five categories describing the number of parents or caregivers and whether or not participants had siblings living in the same household. These categories were collapsed so the data reflected only the number of parents or caregivers in each participants' household. Specifically, data was coded to clearly indicate whether participants did not have a biological or adoptive parent in the home (0), lived in a single parent home (1), or in a two parent home (2).

Reading First status was reported by participants' teachers using three categories: "participates", coded (1), "does not participate", coded (2), and "not offered", coded (3). For each of the variables (kindergarten, first, third, and fifth grades), the data were reverse coded

such that the affirmative, "child participates" responses reflected a higher value (1) than the negative "child does not participate" responses (0). Further, the data originally coded as "not offered" (3) was recoded to "child does not participate" (0) to reflect participants' non-participation in Reading First services in kindergarten (n = 927, 0.09%), first grade (n = 1669, 0.16%), third grade (n = 1226, 11.5%), and fifth grade (n = 1039, 0.10%).

Teacher variables. Across years, teacher education variables consisted of four to five categories ranging from high school education (1) to master's, specialists, or doctoral degree (4) or (5). However, due to small cell sizes, ECLS-K staff collapsed categories with small cell frequencies to protect respondents' confidentiality. As a result, the categories of "high school" and "associate's degree" were combined with "bachelor's degree" into one category representing the highest teacher education levels of high school, an associates degree, and a bachelor's degree. Consequently, it was not possible to evaluate whether the NCLB HQT requirement of "at least a bachelor's degree" (NCLB, 2002) made any effective contribution to sampled participants reading IRT scores.

Therefore, in an effort to evaluate whether teacher educational background contributed to participants reading IRT scores, categories of teacher education were re-coded with regard to whether teachers had earned a graduate degree (1) or any other education level or degree not considered a graduate degree (0). This categorization reflects current discussion in the literature of whether teachers holding a graduate degree can more effectively impact student achievement outcomes (Akbari & Dadvand, 2011; Desimone, 2010; Lewis, 1979; and Staub & Stern, 2002).

Teacher certification status was originally reported over six levels from no certification (0) to advanced certification (5). These categories were collapsed to make the variable a dichotomous choice of "certified" (1) or "not certified" (0) in accordance with the NCLB definition of HQT, which requires that teachers have a bachelor's degree and subject matter knowledge of what they are teaching (NCLB, 2002). With regard to elementary school teachers, any type of teaching certification (provisional, regular, advanced) meets the requirement of subject matter knowledge because of the subject matter contained in the test used to gain certification (NCLB, 2002); categories were collapsed accordingly. Finally, to prepare the teacher degree and teacher certification variables for analysis, they were combined by taking the sum across years to reflect the number of years each participant was taught by a teacher with an advanced degree and/ or teaching certification.

Instruction. All teachers' instructional strategies were recoded dichotomously for MI and regression analysis. While teacher instruction variables were proposed to be combined into composites variables, upon recommendation from the dissertation committee, these variables were reorganized and entered into analysis as independent variables, in groups categorized by cross curricular instruction, choice, and discussion. Table 8 displays the variables used to develop each instructional category, and how the categories were recoded for analysis.

Two instruction variables were measured by duration: time spent on reading or language arts and time spent on child selected activities; the remaining instructional strategies were measured in frequency. Time on reading or language arts was originally coded on a four-point Likert-type scale in 30-minute intervals, ranging from (1) 30 minutes or less; to (4) 90 minutes or more. To better reflect language in NCLB that pertains to funding provided to Reading First

schools who clearly designate 90-minute blocks of time, daily, for reading instruction (NCLB, 2002), the categories were dichotomously collapsed to represent the duration of time of daily reading instruction: (1) 90 minutes or more; or (0) one hour or less. Likewise, a variable that quantified teacher-reported time spent on child-selected activities was originally coded on a 5-point Likert-type scale: (1) no time; (2) 30 minutes; (3) 60 minutes; (4) two hours; and (5) three or more hours. This was recoded to reflect whether teachers spent (1) 60 minutes or more; or (0) 30 minutes or less, on child-selected activities each day.

A measure of participant choice measured the frequency that participants read books of their own choosing during the classroom routine, originally coded (1) almost every day; (2) one to two times per week; (3) one to two times per month; and (4) never or hardly ever. This variable was collapsed into two categories: (1) almost every day; and (0) one to two times per week or less.

Variables used to measure cross-curricular instruction included integrating curricula around common themes and using reading materials drawn from other subjects. Data reflecting whether teachers integrated two or more curriculum areas around common themes was originally coded: (1) never, (2) occasionally; (3) usually or all the time; or (4) all the time. Observations coded as (3) or (4) were recoded (1) usually or all the time; and observations coded (1) or (2) were recoded (0) occasionally or never. Similarly, teachers were asked how often they use reading materials drawn from other subjects, originally coded on a 6-point Likert-type scale: (0) not available; (1) never; (2) one time a month or less; (3) two to three times a month; (4) one to two times per week; (5) three to four times per week ; and (6) daily. These data were collapsed to

reflect: (1) daily participation; or (0) participation 3 to 4 times a week or less, to reflect whether these instructional strategies are firmly embedded in the classroom routine.

The use of discussion strategies in instruction were measured through Likert-type scales that were collapsed and reverse coded to uniformly reflect that higher values indicate positive results. For these variables, the response, (1) almost every day, was not recoded, while the remaining three responses, (2) once or twice a week, (3) once or twice a month, and (4) never or hardly ever, were collapsed into one category, (0) less than every day.

Finally, teachers rated their own feelings of preparation to teach reading and preparation to help with reading problems on a Likert-type scale: strongly disagree (1); disagree (2); ambivalent (3); agree (4); and strongly agree (5). Strongly disagree, disagree, and ambivalent were collapsed into one category labeled, "disagree" (0). Strongly agree and agree were collapsed into one category, labeled, "agree" (1), to clearly distinguish teachers who expressed feelings of adequate preparation to teach reading and assist children who are struggling with reading. Unfortunately, cell frequencies and variability were quite low on teacher preparation variables with most teachers reporting feeling adequately prepared to teach reading and help with reading problems, reducing the likelihood of any significant effect in analyses.

Variable creation. ECLS-K variables measuring the age of participants were originally recorded in units of months. This unit was changed to years, for clarity, by dividing the variable data by 12. Further, the ECLS-K variable reporting the year teachers were born was used to calculate teacher age at the time they completed questionnaires by subtracting the year the teacher was born from the year data was collected. Finally, child retention variables were created for third and fifth grades by subtracting participants reported grade from their expected grade for

third and fifth grade variables to create a variable that reports the number of times a participant had been retained prior to third and fifth grade data collection.

Composite variables. Composite variables were created to better to summarize the extent to which children participated in school programs (e.g., free/reduced lunch, Reading First, special education) and to better summarize participants' characteristics and background at the time of third and fifth grade data collection. As described below, most composite variables used in analysis were created either by taking the sum or the mean of the variables in question.

At each data collection point, participants' school administrators recorded whether their school was public (1) or private (0). The sum of all four variables (kindergarten, first, third, and fifth grades) was used to create a composite variable where participants were coded: (1) "yes" if all four variables were coded "in public school"; and coded (0) "no" if any of the four variables were coded as "not in public school". This variable was used to determine which participants would be excluded from the analytic sample because they were not in public school during all four years of data collection.

On variables describing whether participants changed schools between first, third, and fifth grade data collection, participants were coded: (1) "changed schools" and (0) "did not change schools". The sum of the variables was used to create a cumulative school change variable for third grade (using the sum of first and third grade school change variables) and another for fifth grade (using the sum of first, third, and fifth grade school change variables) to control for any effect of the frequency of school changes on reading achievement. Additionally, dichotomous variables describing involvement in free school lunch, special education and Reading First services were each combined across years by taking the sum of kindergarten, first,

third, and fifth grade variables to reflect the number of years participants were involved with those school activities.

Composite variables including behavioral engagement in third and fifth grade analysis were created by taking the mean score from three (kindergarten, first, third grade) data collection rounds for third grade analysis or four (kindergarten, first, third, fifth grade) data collection rounds for fifth grade analysis. By taking the mean score from three or four data collection rounds (and thus three or four teachers), teacher bias was greatly reduced. Further, the mean score across three or four years was thought to be a strong indicator of children's overall behavioral engagement up to that point in elementary school.

Missing data. As is common in longitudinal survey data, many variables chosen for analysis contained missing values on at least one observation. In this study, 90% of variables (n = 82) in the original data contained missing data and 85% (n = 1594) were complete cases. MI procedures were proposed to recover missing data, and the Statistical Package for the Social Sciences (SPSS) was proposed as the analytical software to accomplish this task. However, due to the complex nature of MI within a large, nested, survey dataset, STATA SE, Version 11.2, was utilized upon the recommendation of a statistical consultant who assisted in these procedures.

***MI preparation*.** Binary variables were originally coded in the ECLS-K dataset with values of (1) indicating an affirmative response and values of (2) indicating a negative response. However, performing MI in STATA SE, requires that binary variables are coded with a range of 0 to 1 (StataCorp., 2009a). Therefore, for each binary variable, negative responses were recoded with values of (0) to uniformly reflect that the higher value of (1) indicates a positive response.

In preparation for MI procedures, Allison (2002) suggests dummy representation of categorical variables, followed by the computation of a reference category as one minus the sum of the dummy categories. Consequently, to prepare the data for MI procedures, all categorical variables were either dummy-coded, or, if appropriate, left alone and treated as continuous variables for the purpose of the MI procedures. Additionally, logarithmic transformations were performed on all continuous variables to keep values within range following MI procedures. The logit of the transformed variables were then prepared to be entered into the MI equation in place of the original variable.

MI procedures. Various options exist within STATA to employ MI procedures. In this research, multivariate normal multiple imputation procedures (MVN) were determined to be the most appropriate due to the presence of multiple variables with missing values and a random pattern of missingness in the data (Allison, 2002; StataCorp, 2009a). Normal distribution is an assumption of univariate MI; however, the MVN model of imputation is robust to situations where some variables may not be normally distributed (Schaefer, 1997). Further, MI using MVN is based on the data augmentation method, producing estimates with little to no bias (Allison, 1999).

Yet, as explained in Stata MI documentation, random patterns of missingness can result in difficulties producing accurate estimates (StataCorp, 2009a). As a result, the Bayesian iterative Markov chain Monte Carlo (MCMC) method was utilized prior to running MVN procedures to ensure that the default, 100 iterations, would be sufficient to approximate the distribution of missing data. The data converged in the MCMC procedure at 78 iterations, suggesting that the default 100 iterations were sufficient to accurately represent the data using MVN MI procedures.

Visual inspection found no pattern in the plot of the worst linear function and a relatively quick decrease in autocorrelation plot, further supporting the decision that the MI MVN procedure would be appropriate in obtaining accurate results. Consequently, MI procedures were completed with 20 additional datasets specified, as recommended in Allison's (2002) work.

Post -MI data management. After examining an overall description of all imputed datasets, visual inspection of summary statistics for each imputed dataset (M = 1/20) were cross-referenced with summary statistics from the original dataset (M = 0). Imputed binary variables were rounded to bring MI values back to binary values of zero or one. Categorical variables, treated as continuous for MI were rounded to bring the imputed values back to integers. Out of range values of continuous variables were brought back within range by exponentiating the logit of the transformed continuous variables to bring the values back between zero and one. Then variables were logarithmically transformed back to their original range of values. Upon visual inspection and comparison with the original data, these data showed no obvious abnormalities and the imputed data were considered an accurate representation of the dataset.

Data Screening

Certain assumptions were necessary to capture accurate results when using multiple regression procedures (Cohen, Cohen, West, & Aiken, 2003). Prior to analysis, variables were screened for outliers and assumptions of multivariate analysis including normality, linearity, homoscedasticity, singularity, multicollinearity, and independence of errors. Univariate box plots and histograms were examined to ensure the assumption of normality was met. Large values of skewness and kurtosis were present in some variables and dealt with on an individual basis, as is explained in further detail below. Bivariate regressions were completed prior to hierarchical

procedures to ensure that each variable introduced into the multivariate equation made a significant contribution to the dependent variable. Only variables that demonstrated a unique individual contribution in bivariate regressions were included in hierarchical analysis.

Univariate and bivariate scatterplots were inspected to rule out any inadvertent effect of outliers and to verify that the assumption of a linear relationship was met. Bivariate correlations between the independent and dependent variables ensured that no multicollinearity or singularity existed between the variables. Further, the variance inflation factor (VIF) and collinearity tolerance were examined for all independent variables included in analyses across three imputed datasets. No issues were found with the variance inflation factor and collinearity tolerance of the independent variables, so these data will not be reported. To ensure that all independent variables demonstrated linearity, normality, and homoscedasticity, residual scatterplots for all independent and dependent variable combinations were visually inspected.

Independent variables were added to the models in a predetermined order, with NCLB (2002) variables (i.e., Reading First intervention and HQT) entered before the process variables (i.e., behavioral engagement and instructional style).The conceptual model for this longitudinal design is presented in Figure 3.

Longitudinal Analyses

Hierarchical regression analyses were employed to test the first two research hypotheses. While multiple regression analyses determine the nature of the relationship between a set of independent variables and a single dependent variable, in this research, independent variables were entered into the hierarchical equation in a specific order to determine what each additional set of variables added to the prediction of the outcome at specific points of entry (Tabachnick &

Fidell, 2001). Hierarchical regression was essential to this study to determine whether the addition of variables not included in NCLB (2002) legislation would improve the prediction of reading outcomes above and beyond what has already been defined in the Act. All analyses were performed with the original (M = 0) data prior to performing analyses on imputed datasets. Visual inspection of results showed no discernible difference between the outcomes, so results from imputed datasets are reported below to ensure that reported results are based on the most complete data for analyses.The relative increase in variance (RVI) was also calculated for third and fifth grade analyses, representing any increases in variance due to missing data. For all regressions, RVI showed very little effect. As there is no standard for RVI, other than a value of zero demonstrating no effect, the small values obtained were deemed satisfactory (Eddings, 2011).

CHAPTER IV: RESULTS

Behavioral Engagement

Hierarchical multiple regression analysis was performed to examine the contribution of children's behavioral engagement skills on their third and fifth grade reading IRT scores to answer the first research question. Participants cumulative behavioral engagement skills were added to the hierarchical regression equations following the inclusion of participant control variables, kindergarten literacy skills and the number of years of participation in Reading First interventions.

Third grade reading scores. Summary statistics of participant control variables are presented in Table 9, and summary statistics for independent variables are presented in Table 10. Displayed in Table 11 are the bivariate coefficients and significance statistics for all variables used to predict third grade reading IRT scores in the first research question. All control and independent variables were significant in third grade bivariate regressions, with the exception of two control variables: participant age at kindergarten entry ($\beta = 0.42$, $p = 0.89$) and participant gender ($\beta = 0.54$, $p = .819$). Consequently, both participant age and gender were not included in the third grade hierarchical equations.

Following bivariate regressions, significant third grade independent variables were entered into the hierarchical regression equation with third grade reading IRT scores as the dependent variable. Table 12 displays the results of each hierarchical regression step. Control

variables that were significant in bivariate regressions were entered in step 1 to account for the effects of children's race, family, and background characteristics. This model was significant ($F_{(14, 332.2)} = 35.16$, $p < .001$), although two variables in the model were not significant, poverty status ($\beta = 3.38$, $p = .126$) and family status ($F_{(2, 265.3)} = .55$, $p = .575$), and consequently they were removed from future models. Individual contributions of the participant control variables entered in the revised step 1 model accounted for 33% of the variance in participants' third grade reading scores[9].

Once participant's kindergarten reading IRT scores were added to the model in step 2, the participants' racial differences ($F_{(6, 336.0)} = 1.94$, $p = .073$) no longer made a statistically significant contribution to their third grade reading scores. Consequently, the participant race variable was removed from future models. The revised model was significant, with SES ($\beta = 3.94$, $p = .004$) and kindergarten reading scores ($\beta = 1.66$, $p < .001$) demonstrating significant positive coefficients and the number of years retained ($\beta = -11.25$, $p < .001$), years in school free lunch program ($\beta = -2.90$, $p < .001$), and years in special education ($\beta = -5.27$, $p < .001$) all demonstrating negative coefficients. In other words, participants from higher SES backgrounds and those with higher kindergarten IRT scores were associated with higher third grade reading IRT scores, and the greater the number of years participants were retained in grade, participated in the free school lunch program, and participated in special education were all associated with lower reading IRT scores in third grade.

When participants' years in Reading First programming was added to the model in step 3, all the variables continued to make significant individual contributions to third grade reading

[9]A caveat of STATA SE 11.2 regression output with MI data is that no R^2 or R^2 change is reported in STATA output. However, programmer's code was obtained from STATA to calculate these statistics for this project.

scores, including Reading First (β = -2.18, p = .014). The details of this significant model are displayed in Table 12, step 3. These data suggest that Reading First participation does contribute to reading outcomes, but not in the way that was expected. Indeed, the data reveal that for each additional year receiving Reading First interventions, participants' scores decreased 2.18 points.

As expected, when participants' cumulative mean behavioral engagement scores (β = 4.96, p = .015) were added in step 4, they contributed above and beyond the contribution of participants' kindergarten reading IRT scores and their cumulative years participating in Reading First programming. The final model containing seven variables is displayed in Table 13, explained 41.6% of the variance in third grade reading IRT scores. This model demonstrates that participants third grade reading scores were 10.14 points lower for each year they were retained in school (range 0 to 3 years), and 4.62 points lower for every year they participated in special education (range 0 to 3 years). A visual inspection of sampled participants' reading scores across years by dichotomous participation (any year/no years), displayed in Figures 4 (special education) and 5 (Reading First) suggests that participation in special education and Reading First services, did not assist in improving participants' scores to the same degree as their peers who did not participate. Indeed, among participants who began school with similar reading IRT scores, an achievement gap developed over time in elementary school between those who did and did not receive services.

Finally, as expected, in the final model, participants's cumulative mean behavioral engagement scores continued to demonstrate a positive contribution to the third grade model above and beyond the contribution of the control variables and participation in Reading First

programming. Specifically, third grade IRT scores (μ = 96.41, *SD* = 22.76, range 45 to 175[10]) increased 4.96 points for every point that participants' third grade cumulative mean behavioral engagement scores (μ = 2.63, *SD* = 0.57, range 1 to 4 scale) increased.

In summary, the final model represented in Table 13, demonstrated that participant SES, grade retainment, free lunch status, years in special education, kindergarten IRT, years in Reading First, and behavioral engagement scores all made a significant contribution to participants' third grade reading IRT scores, confirming the first hypothesis. The final model accounts for over 41% of the variance in children's third grade IRT reading scores, with the independent variables accounting for 8% of the variance beyond that accounted for by the controls.

Fifth grade reading scores. Descriptive statistics of participant independent variables are presented in Table 14. Displayed in Table 15 are the bivariate coefficients and significance statistics for all variables used to predict fifth grade reading IRT scores in the first research question. All variables were significant in fifth grade bivariate regressions, with the exception of two control variables: participant age at kindergarten entry (β = -2.49, p = .402) and participant gender (β= -0.24, p = .905). Consequently, both participant age and gender were not included in the fifth grade hierarchical equations.

Following, the bivariate regressions, significant independent variables were entered into hierarchical regression equations predicting fifth grade reading IRT scores. All control variables that significantly contributed to fifth grade IRT outcome in bivariate regressions were entered in step 1 of the hierarchical regression to control for the effects of participant race, family, and

[10] These data were calculated by taking the mean score of all imputed datasets (n = 20).

background characteristics. Results from this significant model ($F_{(13, 332.2)} = 33.89$, $p < .001$) can be found in step 1 of Table 16. However, three variables in the model, participant poverty in kindergarten ($\beta = 2.86$, $p = .132$), child race ($F_{(6, 333.6)} = 1.77$, $p = .104$) and family status ($F_{(2, 283.8)} = 1.19$, $p = .307$) were not significant, so they were removed from future models prior to entering participants' kindergarten reading IRT scores in step 2. Participant control variables entered in the revised step 1 model accounted for 35% of the variance in participants' fifth grade reading scores.

Participant's kindergarten reading IRT scores were added to the model in step 2, demonstrating a significant contribution over controls, with the overall model demonstrating a significant effect on fifth grade reading IRT scores ($F_{(5, 326.0)} = 103.10$, $p < .001$). Years in Reading First programming was added to the model in step 3 and demonstrated a significant negative contribution similar to that seen with third grade IRT scores with a significant model ($F_{(6, 321.0)} = 85.42$, $p < .001$). As expected, cumulative years in Reading First, contributed above and beyond that of the controls and kindergarten reading scores, but again with a negative coefficient. Finally, in step 4, participants' cumulative mean behavioral engagement scores were added to the model, showing a significant positive effect and a significant overall model ($F_{(7, 311.1)} = 79.87$, $p < .001$).

Table 17 displays the significant final model for fifth grade IRT scores, which demonstrated that, as expected, cumulative mean behavioral engagement scores contributed above and beyond the contribution of control variables, kindergarten reading scores, and participants' cumulative years participating in Reading First reading programming. The seven variables in the final model explained 43.5% of the variance in fifth grade reading IRT scores.

The model shows that participants fifth grade reading scores (μ = 118.75, *SD* =22.71, range 59.63 to 176.49) were over 7.09 points lower for each year they were retained in school (range 0 to 4 years), and 4.62 points lower for every year they were in special education (range 0 to 4 years). Additionally, for each year children spent in Reading First programming (range 0 to 4 years), children's third grade IRT reading scores were 1.78 points lower. However, as participants' behavioral engagement scores (μ = 2.65, *SD* =0.53, range 1 to 4) increased one point, their fifth grade IRT scores increased by 6.57 points.

In summary, similar to the third grade model, the final model showed SES, grade retainment, free lunch status, years in special education, kindergarten reading IRT scores, Reading First participation, and behavioral engagement all demonstrating a significant contribution to participants' fifth grade reading IRT scores. The strongest effect sizes were seen in the number of years retained, number of years in special education, mean behavioral engagement scores, and kindergarten reading IRT scores. The final model demonstrated a moderate effect, accounting for over 43% of the variance in children's fifth grade IRT reading scores with the independent variables accounting for 8% of the variance over that accounted for by the controls.

Teacher Instructional Style

Hierarchical multiple regression analysis was performed to examine the differential contribution of teacher instructional choices on participants' third and fifth grade reading IRT scores to answer the second research question.

Third grade reading scores. Summary statistics for all third grade control and independent variables included in question two analysis can be found displayed in Table 18. Four

teacher variables (teacher certification, elementary certification, preparation to teacher reading, and preparation to help with reading problems) demonstrated a lack of variability and high values of kurtosis; therefore, they were not included in any regression analysis. Prior to beginning the hierarchical regression, bivariate regressions were performed to confirm that each of the control and independent variables made a unique contributions to the dependent variable. Displayed in Table 19 are the bivariate coefficients and significance statistics of the control and independent variables entered with third grade reading IRT scores as the dependent variable. No teacher control or independent variables were significant predictors of third grade reading IRT scores in bivariate regressions. As a result, no further analysis was required to confirm that the model using third grade reading IRT scores as a dependent variable for the second research question was not significant.

Fifth grade reading scores. Summary statistics for all fifth grade control and independent variables included in question two analysis can be found displayed in Table 20. Five teacher variables (teacher certification, elementary certification, discuss interpretation of reading, preparation to teacher reading, and preparation to help with reading problems) demonstrated a lack of variability and high values of kurtosis; therefore, they were not included in any regression analysis.

Prior to beginning the hierarchical regression, bivariate regressions were performed to determine whether each of the control and independent variables made a unique contribution to the dependent variable. Displayed in Table 21 are the bivariate coefficients and significance statistics of the control and independent variables entered with fifth grade reading IRT scores as the dependent variable. Two teacher control variables (teacher race and advanced degree) and

two independent variables (instructional choice and discussion) were significant in fifth grade bivariate regressions. As shown in Table 22, teacher instruction variables were entered into bivariate equations as instructional groups of choice and discussion, following the teacher control variables in step 1. Although, the second research question posited that instructional variables would contribute above and beyond the contribution of teacher certification and instruction time, teacher certification could not be included due to a lack of variance ($\mu = 3.62$, $SD = 0.63$, range 2 to 4), and the instructional time variable ($\mu = 1.45$, $SD = 0.91$, range 0 to 3) was not included in the hierarchical regression because it did not make a significant contribution in bivariate regressions. Therefore, teacher instructional style variables were added to the regression equation following the addition of child and teacher control variables.

Results from the initial model of participant and teacher control variables can be found in step 1 of Table 22. This model demonstrated that although the teacher control variables of teacher race ($\beta = 1.46$, $p = .081$) and teacher advanced degree ($\beta = .74$, $p = .266$) made a significant contribution in bivariate regressions with participants' fifth grade reading scores, they did not significantly contribute above and beyond the contribution of participant control variables of kindergarten SES, cumulative retainment, participation in the school free lunch program, and participation in special education programming. As a result, all teacher control variables were removed from the regression equation prior to entering the first teacher independent variable in step 2. As in question 1, individual contributions of the participant control variables entered in the revised step 1 model accounted for 35.5% of the variance in participants' fifth grade reading scores.

In step 3 of Table 22, the hierarchical regression model examined whether instructional constructs of choice and discussion contributed above and beyond that of the child control variables. The first set of instructional variables to be entered in the hierarchical regression were those describing whether teachers provided participants a choice of child selected instructional activities and/ or a choice of books to read. As individual variables, participation in child selected activities (β = -1.17, p = .214) and providing a choice of books to read (β = 1.39, p = .155) did not add any significant effect to the significant overall model above and beyond that of the participant control variables. Further, when the choice variables were examined as a construct, they did not contribute significantly beyond that of the participant control variables ($F_{(2, 220.1)}$ = 1.88, p = .154). Therefore, the significance of the overall model can only be attributed to participant control variables.

The second set of instructional variables to be entered in the hierarchical regression were those describing whether participants' teachers engaged students in instructional practices involving discussion about what they had read to determine if they contributed to participants fifth grade reading IRT scores above and beyond that of the participant control variables. As a construct, the discussion variables did contribute significantly above and beyond that of participants control variables ($F_{(9, 322.5)}$ = 40.38, p < .001). Yet, among the individual discussion variables, only peer discussion about reading assignments (β = -4.30, p = .011) was found to significantly contribute above and beyond the child control variables. As such, it is likely that this variable is the reason for the significant construct. While instructional constructs of discussion and choice were significant in the model, the effect size was small. Overall, the model

predicting fifth grade reading IRT scores from teacher instructional practices did not confirm the second hypothesis.

Within-Year Analyses

Multiple regression analyses were employed to test the combined contribution of behavioral, instruction, and interaction variables on within-year reading IRT scores. All analyses were performed with the original (M = 0) data prior to performing analyses on imputed datasets. Visual inspection of results showed no discernible difference between the outcomes. Therefore, results from imputed datasets are reported below to ensure that reported results are based on the most complete data for analyses. The RVI was calculated for all within-year analyses, showing very little effect on the increase in variance due to missing data. As with the longitudinal analysis, the small values obtained were considered satisfactory (Eddings, 2011). Prior to beginning the within-year analysis, within year variables of behavioral engagement and teacher instruction were multiplied to create additional within-year behavior-instruction interaction variables.

Kindergarten analyses. Descriptive statistics describing independent variables used in kindergarten analyses can be found in Table 23. Cross-curricular instruction demonstrated a lack of variability, and high values of skewness and kurtosis. Therefore, this variable was not included in regression analysis. Bivariate regressions were performed with all kindergarten independent variables to determine which would be appropriate for inclusion in multiple regression analyses. Table 24 displays the coefficients and significance statistics for all bivariate regressions with kindergarten reading IRT scores as the dependent variable. Bivariate results demonstrated

significant effects for all independent variables, with the exception of choice. Consequently, the choice variable was also left out of the kindergarten within-year multiple regression.

Results from the significant kindergarten multiple regression model ($F_{(6,268.5)} = 10.34$, $p < .001$) can be found in Table 25. This model demonstrates significant main effects of behavioral engagement ($\beta = 3.25$, $p < .001$) and discussion ($\beta = 2.82$, $p < .05$), and no significant interaction effects between kindergarten instructional variables and behavioral engagement beyond individual effects of behavioral engagement and instruction. In other words, the model shows that for each point that participants kindergarten behavioral engagement scores increased (range 1 to 4), their kindergarten reading IRT scores increased by 3.25 points. Further, participants whose kindergarten teachers reported using two types of reading discussion activities regularly in their classrooms had kindergarten reading IRT scores that were 2.83 points higher than those whose teachers only reported using one discussion strategy and 5.66 points higher than those whose kindergarten teachers did not report using any discussion in their reading instruction.

First grade analyses. Descriptive statistics describing independent variables used in first grade analyses can be found in Table 26. Due to high values of kurtosis, the discussion variable was left out of regression analyses. Bivariate regressions were performed with all first grade independent variables to determine which would be appropriate for inclusion in multiple regression analyses. Table 27 displays the coefficients and significance statistics for all bivariate regressions with first grade reading IRT scores as the dependent variable. Bivariate results demonstrated significant effects for behavioral engagement ($\beta = 5.13$, $p < .001$) and interaction effects of behavior-discussion ($\beta = 1.90$, $p = .001$) and behavior-choice ($\beta = 1.06$, $p < .05$).

Consequently, these three independent variables were included in the first grade within-year multiple regression.

Results from the significant first grade multiple regression model ($F_{(3,166.1)} = 6.50$, $p < .001$) can be found in Table 28. This model demonstrates a significant main effect of behavioral engagement ($\beta = 4.39$, $p = .002$), and no significant effects of the interaction between instructional variables and behavioral engagement. In other words, the model shows that for each point that first grade behavioral engagement scores increased (range 1 to 4), first grade reading IRT scores increased by 4.39 points.

Third grade analyses. Descriptive statistics describing independent variables used in third grade analyses can be found in Table 29. All variables were found to be adequate for inclusion in regression analysis. Bivariate regressions were performed with all third grade independent variables to determine which would be appropriate for inclusion in multiple regression analyses. Table 30 displays the coefficients and significance statistics for all bivariate regressions with third grade reading IRT scores as the dependent variable. Bivariate results demonstrated a significant individual effect for behavioral engagement ($\beta = 6.02$, $p < .001$) and no individual instruction or interaction effects. As a result no multiple regression was necessary with third grade within-year analyses. The significant main effect of behavioral engagement demonstrates that for each point that participants third grade behavioral engagement scores increased (range 1 to 4), their third grade reading IRT scores increased by 6.02 points.

Fifth grade analyses. Descriptive statistics describing independent variables used in fifth grade analyses can be found in Table 31. All variables were found to be adequate for inclusion in regression analysis. Bivariate regressions were performed with all fifth grade independent

variables to determine which would be appropriate for inclusion in multiple regression analyses. Table 32 displays the coefficients and significance statistics for all bivariate regressions with fifth grade reading IRT scores as the dependent variable. Bivariate results demonstrated significant effects for behavioral engagement ($\beta = 7.21$, $p < .001$) and significant interaction effects for the variables of behavior/ cross curricular instruction ($\beta = 1.20$, $p = .003$) and behavior/ instructional choice ($\beta = 2.10$, $p = .007$). Consequently, these three independent variables were included in the fifth grade within-year multiple regression.

Significant results for the fifth grade multiple regression model ($F_{(3,255.5)} = 9.26$, $p < .001$) can be found in Table 33. This model demonstrates a significant main effect of behavioral engagement ($\beta = 5.97$, $p < .001$), and no significant effects of the interaction between behavioral engagement-choice or behavioral engagement-cross curricular connections. In other words, the model shows that for each point that fifth grade behavioral engagement scores increased (range 1 to 4), fifth grade reading IRT scores increased by 5.97 points.

Summary of Findings

Behavioral engagement. A main finding of these results is the significant positive longitudinal effect of the behavioral engagement skills of struggling readers on their reading achievement in elementary school, confirming the first hypothesis. A second, related finding is the significant positive within-year effects of behavioral engagement on struggling readers within-year achievement, with the effect size of these results growing as children progressed through elementary school.

School programming. Variables describing participation in school programming efforts of Reading First and special education demonstrated significant negative longitudinal effects,

with effect sizes that were quite large. Further, in-grade retention of struggling readers demonstrated significant negative longitudinal effects.

Child background. As expected, child SES made a significant contribution to the reading achievement of struggling readers in third and fifth grades. Further, child age and kindergarten entry, the number of caregivers living with the child, and child gender were no significant contributors of third or fifth grade reading achievement in this population of struggling readers. Further, child race did not make a significant contribution beyond the effects of kindergarten reading achievement.

Teacher background. Variables describing teacher certification demonstrated low variability and therefore could not be evaluated. Further, variables describing whether teachers held an advanced degree did not make a significant contribution to the reading achievement of struggling readers. In fact, no teacher control or independent variables chosen for analysis made a significant contribution to the third grade reading achievement of struggling readers above and beyond the contribution of child background characteristics. However, teacher race and possession of an advanced degree did make significant individual contributions to the fifth grade reading achievement of struggling readers, but they did not make significant contributions to fifth grade reading scores above and beyond the contribution of child control variables.

Instruction. Few significant longitudinal effects were found with regard to the effect of teacher instructional practices on the reading achievement of struggling readers in third or fifth grades. In fact, no instructional effects were found to significantly contribute to third grade reading achievement in this population. Further, only instructional discussion was found to make a significant contribution to the fifth grade reading achievement of struggling readers, with peer

discussion as the only individual variable that contributed beyond child control variables. Finally, discussion was the only instruction variable to demonstrate a significant contribution in within-year analyses; however, this significant effect was only found in the kindergarten analysis.

CHAPTER V: DISCUSSION

Student behavioral engagement in classroom activities and teacher instructional choices were studied longitudinally within and across school years throughout elementary school. The effect of these variables on student reading achievement was examined with a sample of elementary-aged students who scored at or below the 30th percentile on an assessment of their reading achievement at the end of their kindergarten year. A main finding of this study was that child classroom behavioral engagement significantly contributed to standardized reading scores, both within and across years throughout elementary school. A second finding was that the teacher instructional variables of choice, discussion, and cross curricular instruction did not make consistent significant contributions to student reading outcomes.

Struggling Reader

Discussion of the meaning of these results needs to be made within the context of this sample of struggling readers. All children in this research began formal schooling struggling to perform well on standardized tests of early literacy skills. These children were less likely to recognize letters, beginning sounds, and ending sounds than their same-aged kindergarten peers with standardized scores above the 30th percentile. Additionally, as is consistent with previous research, sampled participants were more likely to present with multiple risk factors for academic difficulties such as race, SES, and the number of parents living in their household (Downer et al., 2007; Livingston & Wirt, 2004; Rathbun et al., 2005; Smokowski et al., 2004;

Vitaro et al., 2005; Zill & West, 2001). Consequently, results from this research can only be generalized to children in this population.

Interestingly, child background characteristics that have previously been associated with lower academic performance in overall samples of children such as race (Snow et al 1998; West et al 2000), age at kindergarten entry (Datar, 2009; Justice et al., 2005), and gender (Denton & West, 2002; Justice et al., 2005; Phillips et al., 2002) did not significantly contribute to reading scores later in elementary school in this population of students, above and beyond the contribution of their kindergarten standardized reading scores. These results are consistent with previous research that some variables commonly associated with early childhood developmental skills lose their significance in early elementary school once the effects of early achievement and school programming are added to the equation.

The current results are consistent with previous evidence suggesting that differences found in the level of kindergarten achievement that have been attributed to entrance age are more likely a reflection of skills learned prior to kindergarten entry than that of child age at school entry (Elder & Lubotsky, 2009). Further, consistent with previous literature, current results demonstrated no significant long-term achievement advantages secondary to age at kindergarten entry for this population of struggling readers (Lincove & Painter, 2006; Stipek & Byler, 2001). Indeed, although Datar (2009) reported a significant positive effect of age at school entry, particularly for children in poverty, using ECLS-K data of an overall sample of first-time kindergarteners, Datar's results did not account for the effect of early achievement or whether the significant age-achievement effects lasted into later elementary school.

Some researchers have discussed parental considerations of resources when making decisions regarding the age their child will enter school (Meisels, 1992 Stipek, 2002). Inherent in this investigation is the suggestion that if age effects do exist, and families from lower SES backgrounds lack the resources to keep their children home or in a childcare setting for an additional year, these children, already at risk, are then at further disadvantage in reaching the achievement level of their peers. However, in the current research, while no significant age effects were found in this sample of struggling readers, lasting significant effects were found for child SES and participation in school free or reduced price lunch. Indeed, child SES, as defined by parent education, parent occupation, and household income level, was positively associated with both third and fifth grade IRT reading scores, and participation in school free or reduced price lunch programming, as an indicator of poverty, was negatively associated with third and fifth grade reading outcome scores (Denton & West, 2002; Sirin, 2005; Snow et al., 1998). The current results are suggestive that focusing on kindergarten entry age for young children at-risk of poor achievement will not yield significant gains in reducing reading achievement gaps.

School programming such as special education and Reading First programs are commonly used as remedial services in schools. Yet, their effectiveness is not well supported in the literature. Achievement gaps within the special education system have been well documented (Chatterji, 2006; Singh, 2013; Zabel & Nigro, 2001), and previous literature has made clear that once children enter special education programming, very few return to regular education (Singh, 2013). Additionally, Gamse et al., (2008) found that increasing instruction time in Reading First programs demonstrated no significant impact on reading comprehension in elementary-aged

students. However, the differential effects of participation in special education among a group of children who all began school struggling with reading skills are not well documented.

Results from this research are consistent with previous literature and add additional information, suggesting that special education and Reading First programs did not provide any advantage to improving reading achievement for this population of students. Indeed, more years of participation in special education and Reading First made a significant negative contribution to third and fifth grade literacy scores. More research is needed in this area to clarify why the gap in achievement widened for children with similar early skill development, specifically examining aspects of special education and Reading First programming that may be more or less effective in improving the reading skill development of struggling readers. These results highlight the importance of using empirically-based reading interventions and instructional strategies, particularly with students who are behind in skill development.

When considering commonly used remedial services in public schools, the effects of retaining children in-grade must be considered. There is much discussion in the literature about the consequences of "social promotion" versus grade retention for struggling students (National Associated of School Psychologists [NASP], 2011). As there is no clear answer for whether to promote children to the next grade when they cannot perform well academically, it is hoped that the current research will add to the evidence used in this discussion. Consistent with previous literature, retaining sampled participants in-grade demonstrated a significant negative contribution to their third and fifth grade reading achievement (Anderson, Whipple, & Jimerson, 2002; Hong & Yu, 2008; Hughes, Kwok, & Im, 2013; Jimerson, Anderson, & Whipple, 2002; Moser, West, & Hughes, 2012; Wu, West, & Hughes, 2008), suggesting that retaining children

in-grade negatively affects achievement outcomes. Yet recently, Im, Hughes, Kwok, Puckett, & Cerda (2013) found no differences in the behavioral engagement or academic achievement of middle schoolers who were differentiated by the presence or absence of elementary school retainment, suggesting that retainment had no effect, positive or negative, on these developmental skills by the time the children were in middle school. Adding to this discussion, Moser and colleagues' (2012) work found initial positive effects of retention that then dissipated before middle school, leaving retained students no better or worse off academically than their non-retained peers.

As this is quite a complicated issue with reciprocal effects of achievement, perceived competence, and socio-emotional and peer repercussions, future research on this topic is necessary. Further, examining the difference in standardized skill development (e.g., ECLS-K proficiency probability scores), behavioral engagement, and behavioral self-ratings of children who were retained compared to those who were not retained might assist in determining whether social-behavioral effects exist, in addition to the academic effects seen with the current sample. Finally, alternatives to retaining children in-grade must be considered and examined with respect to child reading achievement in this population. Only when empirically-validated alternatives to retention are examined in concert with the effect of retention practices on struggling readers will clear answers on this topic begin to emerge.

Behavioral Engagement

The issue of child behavioral engagement is less complicated, both in previous literature and in the current results. As expected, and consistent with previous literature, teacher ratings of child behavioral engagement from kindergarten through third and fifth grades demonstrated a

significant positive effect on sampled participants' third and fifth grade reading scores (Alexander et al., 1993; Claessens et al., 2009; McClelland et al., 2007; McClelland et al., 2000; Ponitz et al., 2009; Spira et al., 2005; Torgeson et al., 1999). This positive effect was significant beyond the individual contribution of kindergarten literacy skills and the combined contribution with participation in Reading First interventions in first and third grade, confirming hypothesis 1. In other words, children who were rated by their teachers as demonstrating better classroom behavioral skills (attention, classroom rule following, eagerness to learn, flexibility, independence, and organization) demonstrated significantly higher reading IRT scores in third and fifth grades. Although McClelland et al., (2006) concluded that kindergarten behavioral engagement skills were not predictive of academic outcomes after second grade (using a general sample of elementary aged children), this research demonstrates that children who are already struggling do show significant effects of behavioral engagement on reading skills beyond the kindergarten year and throughout elementary school.

Current results also demonstrated a significant positive impact of behavioral engagement ratings on within-year reading achievement in kindergarten, first, third, and fifth grades. This finding may be particularly helpful for teachers seeking evidence that improving student behavioral engagement within one school year can raise reading scores. While much of the previous literature on this topic reports results based on a general sample of children (Claessens et al, 2009; DiPerna et al., 2007; McClelland et al., 2007), these results reveal similar results with data focused on children who are in need of additional assistance in reading development, definitively concluding that behavioral engagement is a valuable skill in the elementary school classroom.

The findings of this longitudinal study are the first to confirm the continued importance and positive effect of children's behavioral engagement on reading scores throughout elementary school. Further, these results extended the literature by examining the effectiveness of behavioral engagement above and beyond the contribution of Reading First initiatives. Results demonstrated that not only did behavioral engagement show predictive ability above and beyond Reading First, but more years in Reading First interventions demonstrated a significant negative effect on reading scores. Further, when considering provisions in NCLB referring to increasing time in reading instruction to at least 90 minutes each day (NCLB, 20 U.S.C. §§ 6361-6368 [2002]), the current results did not find a significant effect of increasing instructional time. While NCLB aims to reduce the gap in achievement, increased instructional time in reading and Reading First programming, as prescribed in NCLB, did not demonstrate a significant positive effects on reading achievement for students who begin school with poor reading skill development.

Instruction

When considering characteristics of teacher background and instruction, less definitive results were found in this study of struggling readers. While the work of Rowan et al. (2002) suggested that teacher certification and degree did not demonstrate a significant effect on child reading achievement, and Croninger et al. (2003) found no effects for teacher certification, but positive effects for teacher degree, this study can provide little clarification on this issue. In this study, inconsistent with Croninger et al. (2003), teacher degree did not demonstrate any significant effect beyond that of child control variables. Further, teachers in the study generally held teacher certification and reported using instructional strategies of discussion, choice, and cross-curricular connections. Yet, the variability present on these variables did not significantly

contribute to the prediction of children's third and fifth grade reading scores in this study. As a result, the second hypothesis was rejected. Teachers demonstrated little variability in their instructional practices and their background characteristics were not significant predictors of child achievement beyond child background characteristics. While the lack of variability in teacher certification was unexpected and resulted in the inability to examine the effect of certification status on child reading achievement, this result did provide secondary knowledge that teachers of children in this population tend to hold qualifications identified as HQT.

Further, consistent with Carlisle, Correnti, Phelps & Zeng (2009) teachers' reports of their own content knowledge did not explain child reading achievement. While Stipek (2004) reported that students from low income backgrounds were exposed to less discussion and choice in their academic curriculum, this was not found in the current research. Using an analytic sample of students that was over 35% below the benchmark on federal poverty guidelines, teachers consistently reported daily use of discussion and choice in the reading curricula.

Although Connor et al., (2005) found a significant positive effect of increasing time on instruction and academic performance, the current results do not support the use of increased instructional time to produce higher reading achievement for struggling readers. Notably, Connor et al.'s results were based on a general sample of students, while the current results are based on a sample of children who were struggling with reading skills in kindergarten. It is possible that students who are not struggling with reading skills may benefit from increased reading instructional time, but results from this study suggest that little improvement is gained from instruction time alone for students who are already struggling. Consistent with my results, Block et al. (2009) looked at the reading comprehension in similar-aged students and found that

students given more instructional time without the use of evidence-based instructional styles such as choice, and discussion produced the lowest gains in comprehension.

Previous work in this area was focused primarily on general samples of students, rather than those who are struggling. Using overall samples of students to evaluate the effectiveness of instruction is problematic if the goal is to determine how to improve the scores of children who are struggling in school. The differences between samples may explain some of the inconsistency between previous literature and the current results. While the instructional style constructs of choice, discussion, and cross-curricular instruction previously demonstrated effectiveness in overall samples of students (Block et al., 2009; Mashburn et al., 2008; Parlardy & Rumberger, 2008; Xue & Meisels, 2004), and Hamre and Pianta's (2005) work found gains in student achievement using choice and discussion in the curriculum in a sample of students with characteristics of at-risk backgrounds, this current research did not demonstrate consistent significant effects on the reading scores of this sample of struggling readers. A need exists for research examining the effects of teacher instruction with samples of struggling readers rather than samples of students identified as at-risk using background variables that do not necessarily differentiate those who end up with poorer standardized reading scores.

Finally, consistent with previous literature, longitudinal analysis demonstrated a significant effect of teachers use of classroom discussion on struggling readers fifth grade reading achievement, but no significant effect on third grade longitudinal reading achievement scores. In a comparative study of collaborative peer-managed discussion with teacher managed whole class instruction, Wu, Anderson, Nguyen-Jahiel, and Miller (2013) found low ability students in later elementary school put a higher value on discussion than high ability students.

Yet, in the current study, the influence of peer discussion was not a positive one for struggling readers. The relationship of peer influence on achievement, particularly in low-achieving students, should be explored in future research. As noted in previous literature, the relationship between teacher instruction, classroom behavior, and student achievement is clearly complex and dynamic, with significant reciprocal effects on behavior and achievement (Connor et al., 2011).

Limitations

This research was unique in that it was based on a nationally representative sample of kindergarteners, yet in spite of the large sample of participants, the complex, clustered design of the study resulted in teacher variables being embedded within the child data, leaving interpretation of the teacher data less clear. Further, a lack of variability in teacher reports of certification and feelings of preparation, and small cell sizes on some variables (e.g., teacher degree), left some questions unanswered. Of specific concern were questions regarding the clarity of the HQT status of participants' teachers and whether teachers' instructional practices were effective for this sample of participants.

Further, as this research used secondary data, results are limited to the questions asked in ECLS-K research, and consequently, the lack of information regarding differential curricula in public schools across the country. As discussed in Fredricks et al (2004), many variables not measured in this research could contribute to differences in student achievement, such as the classroom environment, how challenging curricula was for students, the composition of individual classrooms, the effects of peer behavior, and the academic level of peers in classroom. Nuances of teacher's approaches to instruction, teacher-student relationships, peer interactions,

behavioral observations of individual and of individuals within a particular classroom environment all remain unanswered with these data.

Future Research

Only a small number of studies examining child behavior and teacher instruction have focused on early struggling readers, and results from studies using the general population of students cannot be accurately generalized to this population of young children. Exploratory research examining differences between children who begin school with poor pre-reading skills and those who begin with average and above average skills would certainly add to the literature. Further, using a sample of struggling early readers to examine school-based differences between those who improve versus those who do not would add to current knowledge in this area of research. Specifically, there is a need to examine factors beyond child control variables (e.g., race, SES) that may help demonstrate how public schools can intervene and contribute to the improvement of reading scores over time. Future studies using samples of struggling readers to inform remedial reading services is needed, while attempting to differentiate students who improved versus those who did not. Further, future work should examine differences in beginning skills, behavior, school programming, and types of interventions.

Future examination of Reading First services and reading instruction used during extended instruction time in concert with RTI data will better clarify the efficacy of these methods. One examination of RTI reading interventions found support for literacy interventions that emphasized support for engagement and motivation as well as cognitive literacy strategies (Guthrie, Klauda, & Ho, 2013), providing more support for focusing on improving teacher-student relationships in the classroom. Another study suggested that increasing the frequency of

reading tutoring programs was not enough to improve academic outcomes when moving from from tier 2 to tier 3 interventions in RTI programming (Gilbert et al., 2013). Further research is needed, specifically using individualized methods such as RTI in conjunction with in-school graded outcomes and standardized testing outcomes, to determine the most effective combination of behavioral and academic interventions for children who are struggling with reading in elementary school.

In addition, recent trends in educational policy towards standardizing the curricula across the US, often referred to as the Common Core Standards, has highlighted a number of specific literacy skills that children are expected to master by the end of each grade in school (National Governors Association Center for Best Practices & Council of Chief State School Officers [NGACBP], 2010). Future research should examine the impact of the implementation of these policies on child behavior and achievement, investigating concerns over whether these standards may oversimplify the problem of the achievement gap (Bestor, 2014), and potentially interfere with teacher-student relationships. The current research is suggestive that the issue of the achievement gap in the US is a complex, dynamic issue of not only the persons (e.g., teachers and students) and objects (e.g., curricula), but also proximal processes over time (e.g., student classroom engagement). Future research should examine achievement outcomes before and after the implementation of the Common Core Standards with respect to dynamic interactive classroom processes such as behavioral engagement and teacher-student relationships.

Behavioral engagement. Future research should examine the effects of behavioral engagement on standardized reading scores in concert with child grades and self ratings of engagement. While this research study used teacher ratings of behavioral engagement, adding

observation and child self-ratings could clarify behavioral and motivational contributions to school reading achievement (as discussed in Fredricks et al., 2004). Further, measuring tasks of persistence in the classroom (e.g., completion of classroom assignments) separately from externalizing behavioral observations (e.g., rule following) might improve knowledge in research about the differential contributions of these behaviors (as discussed in Fredricks et al., 2004).

A large body of research has examined the impact of children's self-identification as good or poor readers, and the impact of this identity on their choices to engage with texts (Hall, 2009, 2010). Future work using this view of reader self-identity in concert with teacher self-identity (Hall, Johnson, Juzwik, Wortham, & Mosley, 2010) with scales such as The Reader Self Perception Scale (Henk & Melnick, 1995) and behavioral observations of child classroom engagement in reading activities could provide an excellent picture of latent processes occurring within child reading development.

One approach that has shown promise in the literature is the Responsive Classroom Approach (Northeast Foundation for Children, 2007). This approach is based on the premise that meeting children's social and emotional needs in the classroom will assist in improving their academic skills, and emphasizes improving child behavioral skills, referred to in the current study as behavioral engagement skills. Using this approach, researchers have documented improvements in teacher self-efficacy (Rimm-Kaufman & Sawyer, 2004), and in child achievement, social behavior and perceptions of school (Brock, Nishida, Chiong, Grimm, & Rimm-Kaufman, 2008). Further, this approach has been found to contribute to greater gains in child literacy skills, particularly when children participate for more than one year (Rimm-Kaufman, Fan, Chiu, & You, 2007), and produce increased academic gains for children who are

participating in special education services (Elliott, 1993). Most recently, teacher training and adherence to this approach was found to contribute to positive teacher-student relationships (Baroody, Rimm-Kaufman, Larsen, & Curby, 2014). Further empirical validation of teaching methods that assist teachers in combining both academic and social-behavioral classroom goals will be useful in applying this valuable research to practice in US schools.

Additionally, results from the current work, demonstrating that increased time in Reading First programming and special education programming negatively contributed to standardized reading scores merits additional research. The children in this research all began at a similar achievement level on tests of reading skills, yet participation in these programs demonstrated lower reading achievement as participants progressed through school. As special education and Reading First programs are considered remedial services to assist children who are struggling academically, these results provide interesting questions for future study. Certainly, more detailed examination of the differential types of interventions used in Reading First programs and special education instruction might assist in developing some conclusions regarding these gaps.

Finally, this research adds to the literature supporting the trends in research and school policy to minimize or end the use of in-grade retainment (NASP, 2011). Children in this study who were retained demonstrated significantly lower reading achievement with each year of being retained. Future research might look at behavioral engagement differences (observations, child self-report, and teacher ratings) between children who are retained versus those who are not in concert with curricular differences in lower grades.

Instruction. There is a need for longitudinal research examining similar populations of struggling readers who can be grouped differentially by teachers credentials, such as certification or whether they meet the NCLB requirement of "at least a bachelor's degree", thus providing a format to determine whether these requirements are reflective of improved scores for children who are struggling. This current research was unable to provide clarification on this issue.

Future research should also investigate whether lack of teacher variability, definitions of instruction, or differences with an analytic sample of low performing students may have contributed to these results. Due to inconsistent results and the complex nature of teacher instructional variables, future research might further examine the instructional constructs of discussion, choice, and cross-curricular techniques with less stringent categorization than daily use (e.g., two to three times per week).

Further, future examination of other instructional strategies not available in ECLS-K dataset, such as scaffolding techniques (Clark & Graves, 2005) and behavioral aspects of teacher instruction and teacher-student relationships (Baroody et al., 2014; Sableski, 2009) are necessary to differentiate effective instructional strategies from less effective ones for this population of students.

Future research might examine the development of classroom environment measures, especially those that look at proximal processes occurring in the classroom and include teachers as part of the classroom environment (please see Fraser, 1998 for a review of previous known classroom environment scales). Using validated assessment practices designed for implementation with classroom observational techniques, and incorporating the teacher as part of the classroom environment will be imperative in teasing out the nuances of the impact of a

teacher's contribution to child achievement. Two examples of scales are the Classroom Environment Scale (Moos & Tricket, 1987) and the Responsive Environmental Assessment for Classroom Teaching (REACT; Christ, Nelson, & Demers, 2013), though further research is necessary to validated their use in this complex area of research. A third scale that has demonstrated good results in this area is the Classroom Assessment Scoring System (Pianta, La Paro, & Hamre, 2007). Further research examining the effect of classroom and teacher practices on child achievement will be necessary in the development of effective policies and proper appropriation of education funding.

Additionally, as increased instructional time is a provision in NCLB (2002), future research should examine differential instruction that benefits children who are struggling with reading skills to most efficiently use increased instructional time in school, as the effectiveness of increased instructional time is undoubtably linked to the specific types of instructional techniques that are used. Examining teachers instructional strategies more closely to determine what strategies best contribute to improvements in reading scores could be easily researched within an RTI intervention framework. For example, Vernon-Feagons and colleagues (2010) found strong effects on the achievement of struggling readers in kindergarten following a professional development intervention emphasizing tier 2 interventions delivered by the teacher in the classroom environment.

As the knowledge of effective intervention practices increases, research examining the differential effects of teacher professional development is also warranted beyond or instead of concerns about instructional time and/or teacher degree. Of note, recent research has focused on the quality of teacher-student interactions in the classrooms, and the effect that positive (or

negative) interactions may have on students' achievement (Archambault, Pagani, & Fitzpatrick, 2013; Doumen, Koomen, Buyse, Wooters & Verschueren, 2012; Williford, Maier, Downer, Pianta, & Howes, 2013; and Wu, Hughes, & Kwok, 2010). Slavin, Lake, Chambers, Cheung, and Davis (2009) found support for instructional process interventions designed to change behavioral teaching practices and improve teacher-student relationships.

Moreover, Kaiser, Retelsdorf, Sudkamp, and Moller (2013) found significant interactive effects of child behavior and teacher judgements of achievement. Specifically, while teacher-student relationships demonstrate less stability across years than child behavioral engagement ratings (Archambault et al., 2013), responsive and supportive teacher-student relationships have shown strong within-year effects on student achievement (Doumen et al., 2012; Williford et al., 2013; Wu et al., 2010). Further, Ponitz et al. (2009) described the interaction between instruction and behavior, looking at overall classroom quality, measured through the use of classroom observation and including the constructs of student emotional support and classroom organization. The interaction seen in Ponitz's work may reflect characteristics, such as the quality of teacher-student interactions that are best measured using classroom observation. While teacher instructional style may not have had significant effects on child achievement, observations of teacher behavior during instruction and reports of teacher-student interactions during classroom activities may have a significant effect on student achievement.

Finally, Kennedy (2010) reported good results with teacher professional development focused on classroom management techniques, literacy instruction, strategies using higher order thinking skills and using meaningful contexts to teach basic skills. While the sense of preparation by teachers was not significant in this research, examining completed coursework and

implementation of professional development opportunities that focus on evidence-based interventions might demonstrate more convincing results in improving the skill level of struggling readers.

Potential Applications of Findings

The findings of this study support the current trends in public education of the early identification and treatment of behavioral and academic difficulties using individualized assessment and intervention procedures such as RTI, over the use of general school programming such as increased instructional time, special education, grade retention, and Reading First programs. RTI provides the benefit of individualized programming with progress monitoring to continuously evaluate and improve the effectiveness of interventions. Continued use and strengthening of RTI programs in school systems should decrease the use of special education placement and provide better alternatives to retention for young children. Further, as part of RTI systems, children's behavioral engagement should be monitored and treated to improve school engagement and child achievement. Finally, it is recommended that teacher professional development sessions should be focused on strategies to improve behavioral engagement in the classroom, and improving teacher skills using RTI for academic and behavioral progress monitoring and treatment effectiveness.

A final contribution of the findings of this study is to inform education policy efforts, such as that seen in NCLB (2002). In this regard, constructs highlighted in NCLB such as teacher education, Reading First programs, and increased instructional time appear to be less effective than desired at producing the necessary improvements in standardized reading scores for students who begin school struggling with early reading skills. Findings of this research, in conjunction

with future efforts, offer information for policymakers to develop evidence-based educational policies and improve the quality of public education for our children.

Conclusion

Child behavioral engagement in classroom activities is an important component of child reading achievement that has been overlooked in federal education regulations. Research is increasing support for developing child behavioral engagement skills as a part of academic curricula, and future efforts to bridge this research-to-policy gap may prove quite successful in reducing gaps of reading achievement that persist in public schools. Second, evaluating teachers as "Highly Qualified" based on their certification and highest level of education, as is current practice, has not demonstrated conclusive positive results in terms of student gains in research efforts. Future efforts should continue to examine teacher-student processes that may assist in improve reading gains for struggling readers in our public schools.

Table 1.

30th Percentile Kindergarten IRT Reading Scores Across Imputed Datasets and Computed Mean IRT Reading Score Used to Create Sample (n = 10651)

Imputed Dataset	30th percentile IRT score	95% *CI*
M = 1	34.07	[33.83, 34.33]
M = 2	34.13	[33.89, 34.39]
M = 3	34.05	[33.81, 34.30]
M = 4	34.08	[33.84, 34.34]
M = 5	34.04	[33.81, 34.28]
M = 6	34.06	[33.81, 34.30]
M = 7	34.04	[33.81, 34.30]
M = 8	34.11	[33.87, 34.38]
M = 9	34.08	[33.84, 34.34]
M = 10	34.06	[33.81, 34.32]
M = 11	34.06	[33.81, 34.31]
M = 12	34.01	[33.79, 34.25]
M = 13	34.03	[33.80, 34.27]
M = 14	34.10	[33.87, 34.36]
M = 15	34.02	[33.80, 34.26]
M = 16	34.04	[33.81, 34.28]
M = 17	34.06	[33.82, 34.30]
M = 18	34.10	[33.88, 34.37]
M = 19	34.06	[33.81, 34.32]
M = 20	34.04	[33.81, 34.30]
μ M(1, 20)	34.062	

Note. CI = confidence interval.

Table 2.

Comparison of Sampled and Excluded Participants' Demographic Characteristics

	Sampled n = 1880	Excluded n = 8793	Significance (sample - excluded)	
Child Variables	%	%	*t* value	*p* value
Below poverty level	**35.24**	**18.94**	**7.47**	**< .001**
Family Status				
No biological or adoptive parents	**2.74**	**1.18**	**2.07**	**.039**
One parent household	**37.61**	**20.27**	**7.42**	**< .001**
Two parent household	**59.64**	**78.56**	**-7.99**	**< .001**
Free lunch				
Kindergarten	**47.92**	**30.48**	**6.74**	**< .001**
First grade	**60.46**	**35.19**	**10.73**	**< .001**
Third grade	**60.06**	**35.95**	**10.23**	**< .001**
Fifth grade	**60.1**	**35.00**	**9.88**	**< .001**
Gender				
Female	**42.60**	**50.24**	**-3.12**	**.002**
Male	**57.40**	**49.76**	**3.12**	**.002**
Race				
Asian	**1.13**	**3.32**	**-4.75**	**< .001**
Black	**30.02**	**12.41**	**6.91**	**< .001**
Hispanic	**12.65**	**20.47**	**-4.88**	**< .001**
Multiracial	2.07	2.35	-0.54	.571
Native American	3.38	1.28	1.40	.163
Pacific Islander	0.74	0.51	0.79	.430
White	**50.01**	**59.66**	**-3.39**	**.001**
Region: USA				
Northeast	17.54	17.83	-0.13	.899
South	**43.39**	**35.14**	**2.78**	**.006**

	Sampled	Excluded	Significance	
	n = 1880	n = 8793	(sample - excluded)	
Child Variables	%	%	*t* value	*p* value
Midwest	24.17	23.17	0.31	.754
West	**14.80**	**23.87**	**-4.72**	**< .001**
Region: Local				
City	35.20	37.23	-0.65	.517
Suburb	38.71	42.51	-1.11	.267
Rural	**25.99**	**20.27**	**2.03**	**.043**
Retained				
First grade	**8.38**	**2.5**	**5.06**	**< .001**
Third grade	**27.19**	**5.94**	**9.67**	**< .001**
Fifth grade	**32.26**	**8.58**	**9.31**	**< .001**
Special Education				
Kindergarten	**9.24**	**3.25**	**3.85**	**< .001**
First grade	**10.91**	**3.41**	**5.51**	**< .001**
Third grade	**18.41**	**6.37**	**6.36**	**< .001**
Fifth grade	**23.04**	**7.64**	**7.62**	**< .001**
Reading First				
Kindergarten	**17.78**	**8.64**	**4.46**	**< .001**
First grade	**34.73**	**17.26**	**6.64**	**< .001**
Third grade	**26.06**	**14.72**	**5.30**	**< .001**
Fifth grade	**20.14**	**9.63**	**4.83**	**< .001**

Table 3.

Comparison of Percentage of Sampled and Excluded Participants' Hierarchical Mastery[11] of Early Literacy Skill Development at Kindergarten Entry

Hierarchical Skill Level	% Participants		*p* value
	Sampled (n = 1880)	Excluded (n = 8793)	
< Letter recognition	20.9	3.8	< .001
Letter recognition	52.5	11.1	< .001
Beginning sounds	25.0	25.0	.980
Ending sounds	1.6	43.8	< .001
Sight words	0	11.6	
Total percent	100	95.3	

[11] Mastery of one level assumes mastery of previous levels and non-mastery of subsequent levels (Tourangeau, 2006)

Table 4.

Frequency of Sampled Participants' Teachers Who Completed Multiple Participant Questionnaires

Frequency of Completed Questionnaires	Frequency of Teachers by Grade					
	Kindergarten		Third		Fifth	
	n	(%)	n	(%)	n	(%)
1	599	(37)	747	(55)	928	(57)
2	372	(23)	346	(25)	350	(22)
3	249	(16)	141	(10)	153	(10)
4	140	(9)	24	(2)	60	(4)
5	60	(4)	40	(3)	35	(2)
6	42	(3)	30	(2)	42	(3)
7	28	(2)	0	(0)	0	(0)
8	40	(2)	16	(1)	16	(1)
9	9	(1)	18	(1)	9	(1)
10	40	(2)	0	(0)	0	(0)
11	11	(1)	0	(0)	11	(1)
14	14	(1)	0	(0)	0	(0)

Note. These data are based on un-imputed estimates (M = 0) and are rounded to the nearest whole number.

Table 5.

Descriptive Statistics for Sampled Participants' Teachers Across Grade Levels

	Participant Grade			
	Kindergarten	First	Third	Fifth
Advanced degree (%)	35.8	45.5	42.9	44.5
Certification (%)	91.3	91.3	89.9	88.6
Elementary education certification (%)	n/a	93.0	93.6	94.2
Mean age (years)	41.65	43.13	42.3	43.2
Years teaching (years)	n/a	15.33	14.95	14.2
Race (%)				
Asian/ Native American/ Pacific Islander	3.1	2.7	n/a	n/a
Black	10.5	9.9	n/a	n/a
Hispanic	1.3	0.4	n/a	n/a
Multiracial	2.0	3.0	n/a	n/a
Nonwhite	n/a	n/a	16.5	18.6
White	84.1	84	83.5	81.4

Note: To ensure confidentiality of participants, some racial categories were collapsed by ECLS-K staff due to small cell size (Tourangeau, 2006).

Table 6.

Summary Statistics of Sampled Participants Scores (Range 1, 4) on Individual Measures of Behavioral Engagement in Kindergarten, First, Third, and Fifth Grades using un-imputed data (M=0)

Measures of Behavioral Engagement	*μ (SD)*	Skewness	Kurtosis
Kindergarten			
Attention (n = 1566)	2.66 (.86)	.12	2.17
Eagerness (n = 1566)	2.87 (.86)	-.13	2.06
Flexibility (n = 1563)	2.86 (.81)	-.15	2.29
Independence (n =1568)	2.76 (.86)	.01	2.09
Organization (n = 1552)	2.72 (.90)	-.10	2.15
Persistence (n = 1566)	2.78 (.90)	-.15	2.14
First Grade			
Attention (n = 1464)	2.60 (.88)	.17	2.19
Eagerness (n = 1467)	2.83 (.85)	-.06	2.08
Flexibility (n = 1462)	2.78 (.81)	.003	2.24
Independence (n =1467)	2.69 (.85)	.06	2.20
Organization (n = 1467)	2.60 (.96)	.01	2.01
Persistence (n = 1461)	2.69 (.92)	-.03	2.07
Third Grade			
Attention (n = 1355)	2.57(.87)	.22	2.25
Classroom rule following (n =1359)	3.05(.80)	-.23	1.96
Eagerness (n = 1356)	2.57(.95)	-.01	2.05
Flexibility (n = 1349)	2.77(.81)	-.13	2.41
Independence (n = 1360)	2.72(.86)	.12	2.05
Organization (n = 1356)	2.73(.85)	.09	2.09
Persistence (n = 1359)	2.67(.89)	.001	2.17
Fifth Grade			
Attention (n = 1545)	2.63(.83)	.14	2.28
Classroom rule following (n = 1547)	3.15(.78)	-.35	2.01
Eagerness (n = 1543)	2.65(.82)	.19	2.25
Flexibility (n = 1534)	2.83(.79)	-.09	2.35
Independence (n = 1547)	2.82(.82)	-.03	2.15
Organization (n = 1535)	2.66(.93)	-.07	2.09
Persistence (n = 1542)	2.73(.89)	-.05	2.13

Table 7.

Frequency and Explanation and Recoding of "Not Applicable" Missing Values for all Eligible Participants (n = 10,673)

	Not Applicable Missing Values					
	Frequency in Grade					
Variable	K	1	3	5	Explanation	Recoded Value
Choice of books to read	--	--	124	--	non-response	system missing (.)
Difficulty hearing	9,174	--	--	--	no hearing problem	no (0)
Difficulty vision	8899	--	--	--	no vision problem	no (0)
Discuss interpretation of what read	--	--	124	--	non-response	system missing (.)
Discuss new vocabulary	--	--	124	--	non-response	system missing (.)
Explain understanding of what read	--	--	124	--	non-response	system missing (.)
Free lunch	4644	2391	2284	2265	does not participate	no (0)
Group activity/ project about book read	--	--	124	--	non-response	system missing (.)
School change	--	1	17	29	homeschooled	no (0)
Talk about book read	--	--	124	--	non-response	system missing (.)
Teaching certification	--	--	--	128	no certification	none (0)
Time reading	1	--	111	4	topic not covered	90 minutes or fewer (0)

Note: Homeschooled children were recoded to No (0) because they did not meet the criteria of moving between two public schools.

Table 8.

Categorization of Instructional Variables for Analysis.

Instructional Construct	Variable	Original Categories	Recoded Categories
Choice			
	Read books they have chosen for themselves	1 = almost every day 2 = once or twice per week 3 = once or twice per month 4 = never or hardly ever	0 = one to two times per week or less 1 = almost every day
	On a typical day, how much time to the children spend in child-selected activities	1 = no time 2 = half hour or less 3 = about one hour 4 = about two hours 5 = three hours or more	0 = 30 minutes or less 1 = 60 minutes or more
Cross-curricular			
	To what extent do you integrate curriculum around common or unifying themes	1 = never 2 = occasionally 3 = usually 4 = all the time	0 = occasionally or never 1 = usually or all the time
	How often do your children use reading materials drawn from other subject areas	0 = not available 1 = never 2 = once a month or less 3 = two or three times a month 4 = once or twice a week 5 = three or four times a week 6 = daily	0 = participation three to four times a week or less 1 = daily participation
Discussion			
	Discuss new or difficult vocabulary	1 = almost every day 2 = once or twice per week 3 = once or twice per month 4 = never or hardly ever	0 = less than every day 1 = almost every day
	Talk with others about what they are reading	1 = almost every day 2 = once or twice per week 3 = once or twice per month 4 = never or hardly ever	0 = less than every day 1 = almost every day
	Do a group activity or project about what they have read	1 = almost every day 2 = once or twice per week 3 = once or twice per month 4 = never or hardly ever	0 = less than every day 1 = almost every day

Instructional Construct	Variable	Original Categories	Recoded Categories
	Discuss different interpretations about what they have read	1 = almost every day 2 = once or twice per week 3 = once or twice per month 4 = never or hardly ever	0 = less than every day 1 = almost every day
	Explain or support their understanding about what they have read	1 = almost every day 2 = once or twice per week 3 = once or twice per month 4 = never or hardly ever	0 = less than every day 1 = almost every day
Instruction time	How much time do children in your classes usually work on lessons or projects in reading and language arts	1 = 1 - 30 minutes a day 2 = 31 - 60 minutes a day 3 = 61 - 90 minutes a day 4 = More than 90 minutes a day	0 = one hour or less 1 = 90 minutes or more
Teacher preparation			
	I am adequately prepared to teach reading to the children who are in my class	1 = strongly disagree 2 = disagree 3 = neither agree nor disagree 4 = agree 5= strongly agree	0 = disagree 1 = agree
	I am adequately prepared to assist children who are experiencing difficulties in reading	1 = strongly disagree 2 = disagree 3 = neither agree nor disagree 4 = agree 5= strongly agree	0 = disagree 1 = agree

Table 9.

Descriptive Statistics of Third Grade Control Variables for Sampled Participants in the First Research Question (n = 1880)

Variable	μ	*SD*	Range	Skewness	Kurtosis
Age (years)	6.14	0.35	5.37, 7.71	0.49	3.19
Gender (%) (0 =female; 1 = male)	0 = 42.6 1 = 57.4	0.49	0, 100	-0.33	1.1
Kindergarten poverty (%) (0= below; 1 = at or above)	0 = 35.2 1 = 64.8	0.45	0, 100	-0.94	1.88
Kindergarten SES	-0.27	0.70	-4.47, 2.62	-0.55	7.66
Family	%				
No parents	2.7	0.17	0, 100	5.58	32.08
One parent	37.6	0.46	0, 100	0.88	1.77
Two parents	59.6	0.47	0, 100	-0.73	1.53
Race	%				
Asian	1.1	0.15	0, 100	6.37	41.58
Black	30.0	0.43	0, 100	1.21	2.47
Hispanic	12.6	0.33	0, 100	2.31	6.32
Multiracial	2.1	0.16	0, 100	6.00	37.05
Native American	3.4	0.19	0, 100	4.84	24.42
Pacific Islander	0.7	0.12	0, 100	7.85	62.6
White	50.0	0.50	0, 100	-0.14	1.01

Note: These data were calculated by taking the mean of each variable across imputed datasets (n = 20).

Table 10.

Descriptive Statistics of Third Grade Independent Variables for Sampled Participants in the First Research Question (n = 1880)

Variable	μ	Range	*SD*	Skewness	Kurtosis
Behavior engagement	2.63	1, 4	0.57	0.02	2.36
K IRT	28.67	16.6, 34.04	3.84	- 0.78	2.91
Third IRT	96.41	45.57, 174.8	22.76	0.04	2.73
Years in free lunch	1.68	0, 3	1.29	0.06	1.29
Years retained cumulative	0.28	0, 3	0.43	1.55	4.34
Years in special education	0.39	0, 3	0.75	2.02	6.26
Years in Reading First	0.79	0, 3	0.92	0.86	2.69

Note: These data were calculated by taking the mean of each variable across imputed datasets (n = 20).

Table 11.

Individual Contribution of Independent Variables in Bivariate Regressions with Third Grade IRT Reading Scores as the Dependent Variable (n = 1880)

Variables	*β*	*SE*	*t* score	*95% CI*	*F (df)*	*p* value
Age kindergarten entry	0.42	3.01	0.14	-5.50, 6.34	0.02 (1, 328.0)	.890
Gender	0.54	2.36	0.23	-4.10, 5.18	0.05 (1, 330.8)	.819
Race					**10.17 (6, 336.5)**	**< .001**
Black	-10.76	2.23	-4.83	-15.14, -6.38		< .001
Hispanic	-10.43	3.03	-3.44	-16.40, -4.47		.001
Asian	-4.73	4.61	-1.03	-13.79, 4.34		.306
Pacific Islander	0.53	6.28	0.08	-11.83, 12.89		.933
Native American	**-19.28**	**3.72**	**-5.19**	**-26.59, -11.98**		**< .001**
Multiracial	-6.57	4.64	-1.42	-15.71, 2.56		.158
SES	**9.83**	**1.58**	**6.24**	**6.73, 12.93**	**38.92 (1, 284.4)**	**< .001**
K poverty status	**12.58**	**2.20**	**5.71**	**8.24, 16.92**	**32.62 (1, 210.6)**	**< .001**
Free lunch	**-6.22**	**0.72**	**-8.61**	**-7.65, -4.80**	**74.08 (1, 310.0)**	**< .001**
Years in special education	**-9.13**	**1.08**	**-8.49**	**-11.25, -7.01**	**72.04 (1, 290.7)**	**< .001**
Retained	**-18.94**	**1.88**	**10.07**	**-22.65, -15.24**	**101.38 (1, 321.9)**	**< .001**
Family status					**4.80 (2, 282.2)**	**.009**
No parent in home	-5.84	5.94	-0.98	-17.53, 5.85		.326
One parent in home	**-6.08**	**2.16**	**-2.81**	**-10.35, -1.81**		**.006**
K IRT	**2.70**	**0.21**	**12.87**	**2.28, 3.11**	**165.59 (1, 311.8)**	**< .001**
Behavioral engagement	**13.34**	**2.23**	**5.99**	**8.91, 17.76**	**35.86 (1, 92.4)**	**< .001**
Reading First participation	**-4.58**	**1.01**	**-4.54**	**-6.58, -2.59**	**20.63 (1, 135.2)**	**< .001**

Table 12.

Hierarchical Multiple Regression Analyses Predicting Third Grade IRT Reading Scores from Kindergarten Reading IRT Scores, Behavioral Engagement Scores, and Reading First Status

		ΔR	β	*SE*	95% *CI*	*F (df)*	*p* value
Step 1		0.34				**35.16 (14, 332.2)**	**< .001**
	Race					**3.12 (6, 335.6)**	**.006**
	SES		**3.47**	**1.53**	**0.45, 6.48**		**.024**
	Poverty		3.38	2.20	-0.96, 7.72		.126
	Retained		**-14.48**	**1.88**	**-18.17, -10.78**		**< .001**
	Free lunch		**-2.52**	**0.96**	**-4.41, -0.63**		**.009**
	Years special education		**-7.78**	**1.02**	**-9.80, -5.78**		**< .001**
	Family					0.55 (2, 265.3)	.575
Step 2		0.07				**44.15 (11, 335.1)**	**< .001**
	K IRT		**1.61**	**0.23**	**1.16, 2.05**		**< .001**
Step 3		0.008				**84.31 (6, 308.6)**	**< .001**
	Reading First		**-2.18**	**0.88**	**-3.91, -0.45**		**.014**
Step 4		0.014				**71.28 (7, 312.4)**	**< .001**
	Behavioral engagement		**4.96**	**2.00**	**0.98, 8.94**		**.015**

Note: The final model is displayed in Table 13.

Table 13.

Final Model of Hierarchical Multiple Regression Analyses Predicting IRT Reading Scores in Third Grade from Kindergarten Reading IRT Scores, Behavioral Engagement Scores, and Reading First Status

	Third Grade IRT Reading Scores				
Predictor	*β*	*SE*	*t score*	*95% CI*	*p* value
SES	**3.79**	**1.37**	**2.77**	**1.09, 6.48**	**.006**
Retained	**-10.14**	**1.74**	**-5.83**	**-13.56, -6.71**	**< .001**
Years in free lunch	**-2.58**	**0.83**	**-3.10**	**-4.21, -0.94**	**.002**
Years in special education	**-4.62**	**1.01**	**-4.59**	**-6.60, -2.64**	**< .001**
K IRT reading	**1.51**	**0.23**	**6.62**	**1.06, 1.97**	**< .001**
Child behavior	**4.96**	**2.00**	**2.47**	**0.98, 8.94**	**.015**
Years in Reading First	**-2.03**	**0.88**	**-2.32**	**-3.76, -0.30**	**.027**
Final Model	$R^2 = 41.65$				**< .001**

Table 14.

Descriptive Statistics of Fifth Grade Independent Variables used in the First Research Question (n = 1880)

Variable	μ	Range	*SD*	Skewness	Kurtosis
Behavioral engagement	2.65	1, 4	0.53	0.01	2.32
Fifth IRT	118.75	59.63, 176.49	22.71	-0.28	2.76
Years in free lunch	2.29	0, 4	1.70	-0.01	1.30
Years retained, cumulative	1.35	0, 4	0.48	1.43	4.27
Years in special education	0.62	0, 4	1.04	1.65	4.71
Years in Reading First	0.99	0, 4	1.07	0.88	2.91

Note: These data were calculated by taking the mean of each variable across imputed datasets (M = 20).

Table 15.

Individual Contribution of Independent Variables in Bivariate Regressions with Fifth Grade IRT Reading Scores (n = 1880)

Variables	*β*	*SE*	*t* score	*95% CI*	*F (df)*	*p* value
Age kindergarten entry	-2.49	2.97	-0.84	-8.34, 3.36	0.70 (1, 329.8)	.402
Gender	-0.24	2.04	-0.12	-4.25, 3.77	0.01 (1, 334.0)	.905
Race					**6.53 (6, 336.7)**	**< .001**
Black	-11.58	2.51	-4.61	16.52, -6.64		
Hispanic	-8.48	2.98	-2.85	-14.34, -2.62		
Asian	0.68	5.11	0.13	-9.37, 10.74		
Pacific Islander	-2.25	5.65	-0.40	-13.38, 8.88		
Native American	-16.96	4.56	-3.72	-25.93, -8.00		
Multiracial	-6.37	4.89	-1.30	-15.98, 3.25		
SES	**11.34**	**1.20**	**9.47**	**8.98, 13.69**	**89.72 (1, 297.8)**	**< .001**
K poverty status	**12.94**	**1.74**	**7.42**	**9.51, 16.38**	**55.03 (1, 218.3)**	**< .001**
Years in free lunch	**-5.12**	**0.48**	**-10.74**	**-6.06, -4.18**	**115.27 (1, 308.9)**	**< .001**
Years in special education	**-7.85**	**0.89**	**-8.87**	**-9.60, -6.11**	**78.60 (1, 269.0)**	**< .001**
Years retained, cumulative	**-15.11**	**1.95**	**-7.76**	**-18.94, -11.28**	**60.20 (1, 331.4)**	**< .001**
Family status					**4.96 (2, 316.5)**	**.008**
No parent in home	-6.42	6.23	-1.03	-18.68, 5.83		
One parent in home	-5.80	2.12	-2.73	-9.98, -1.62		
K IRT	**2.63**	**0.20**	**13.01**	**2.23, 3.03**	**169.26 (1, 255.6)**	**< .001**
Behavioral engagement	**15.98**	**1.94**	**8.23**	**12.14, 19.83**	**67.74 (1, 114.1)**	**< .001**
Reading First participation	**-4.42**	**0.87**	**-5.05**	**-6.14, -2.69**	**25.55 (1, 186.5)**	**< .001**

Table 16.

Hierarchical Multiple Regression Analyses Predicting Fifth Grade IRT Reading Scores from Kindergarten Reading IRT Scores, Behavioral Engagement Scores, and Reading First Status

	Δ*R*	*β*	*SE*	95% *CI*	*F (df)*	*p* value
Step 1	0.37				**33.89 (13, 332.2)**	**< .001**
Race					1.77 (6, 333.6)	.104
SES		**4.94**	**1.31**	**2.37, 7.52**		**< .001**
Poverty		2.86	1.89	-0.87, 6.58		.132
Years retained, cumulative		**-11.00**	**2.01**	**-14.96, -7.03**		**< .001**
Years in free lunch		**-2.09**	**0.62**	**-3.31, -0.87**		**.001**
Years in special education		**-7.27**	**0.85**	**-8.95, -5.60**		**< .001**
Family status					1.19 (2, 283.8)	.307
Step 2	0.06				**103.10 (5, 326.0)**	**< .001**
K IRT		**1.48**	**0.20**	**1.08, 1.88**		**< .001**
Step 3	0.01				**85.42 (6, 321.0)**	**< .001**
Years in Reading First		**-1.89**	**0.59**	**-3.05, -0.72**		**.002**
Step 4	0.02				**79.87 (7, 311.1)**	**< .001**
Behavioral engagement		**6.57**	**1.65**	**3.29, 9.85**		**< .001**

Note: The final model is displayed in Table 17.

Table 17.

Final Model of Hierarchical Multiple Regression Analyses Predicting IRT Reading Scores in Fifth Grade from Kindergarten Reading IRT Scores, Behavioral Engagement Scores, and Reading First Status

	Fifth Grade IRT Reading Scores				
Predictor	*β*	*SE*	*t* score	*95% CI*	*p* value
SES	**5.05**	**1.27**	**3.99**	**2.55, 7.55**	**< .001**
Retained	**-7.09**	**1.66**	**-4.26**	**-10.36, -3.82**	**< .001**
Years in free lunch	**-2.11**	**0.56**	**-3.78**	**-3.21, -1.01**	**< .001**
Years in special education	**-4.62**	**0.85**	**-5.41**	**-6.30, -2.94**	**< .001**
K IRT reading	**1.33**	**0.20**	**6.63**	**0.94, 1.73**	**< .001**
Behavioral engagement	**6.57**	**1.65**	**3.97**	**3.29, 9.85**	**< .001**
Years in Reading First	**-1.78**	**0.60**	**-3.00**	**-2.96, -0.61**	**.003**
Final Model	R^2				**< .001**

Table 18.

Summary Statistics of Third Grade Independent Variables used in the Second Research Question (n = 1880)

Variable	μ	*SD*	Range	Skewness	Kurtosis
Teacher age	42.39	6.65	25, 61	-0.11	2.70
Years with certified teacher	2.73	0.62	1, 3	-1.45	3.94
Elementary certification	1.88	0.28	1, 2	-2.93	9.59
Advanced degree	1.26	0.98	0, 3	0.29	2.06
Years teaching	15.13	7.86	1, 35	0.27	2.35
Choice					
Frequency child chooses books	2.46	0.67	1, 3	-0.74	2.42
Time on child-selected activities	1.14	0.85	0, 3	0.28	2.37
Cross-curricular connections					
Integrate curriculum around themes	2.17	0.71	1, 3	-0.22	2.00
Read other subjects daily	0.54	0.50	0, 1	-0.07	1.00
Discussion (frequency)					
Discuss interpretation of reading	0.25	0.43	0, 1	1.18	2.40
Discuss new vocabulary	2.31	0.70	1, 3	-0.48	2.10
Group projects about a book	0.75	0.71	0, 2	0.37	2.01
Child explain understanding	0.44	0.50	0, 1	0.22	1.05
Child talks about what read	0.45	0.50	0, 1	0.19	1.04
Time on reading instruction	1.22	0.82	0, 3	0.22	2.49
Preparation to teach	0.95	0.23	0, 1	-3.84	15.73
Preparation to help	0.85	0.36	0, 1	-1.92	4.69

Note: These data were calculated by taking the mean of each variable across imputed datasets (M = 20).

Table 19.

Individual Contribution of Teacher Independent Variables used in the Second Research Question in Bivariate Regressions with Third Grade Reading IRT Scores (n = 1880)

Variables	*β*	*SE*	*95% CI*	*t* score	*p* value
Teacher age	0.09	0.14	-0.17, 0.36	0.69	.489
Teacher race	2.45	1.36	-0.22, 5.12	1.81	.072
Years teaching	0.13	0.13	-0.13, 0.40	1.01	.314
Advanced degree	1.42	1.29	-1.12, 3.95	1.10	.272
Time on reading instruction	-0.61	1.36	-3.30, 2.09	-0.45	.657
Choice		F (2, 240.3) = 1.40			.248
Frequency child chooses books to read	0.31	1.76	-3.17, 3.78	0.17	.862
Time on child selected activities	-2.08	1.28	04.61, 0.45	-1.62	.107
Cross-curricular connections		F (2, 256.9) = 0.06			.938
Integrate curriculum around themes	-0.50	1.46	-3.38, 2.38	-0.34	.732
Read other subjects daily	0.21	2.39	-4.49, 4.92	0.09	.930
Discussion (frequency)		F (5, 285.6) = 0.25			.938
Discuss new vocabulary	-1.39	1.63	-4.60, 1.82	-0.85	.394
Group projects about a book	-0.13	1.41	-2.92, 2.66	-0.09	.929
Discuss interpretation of reading	-0.12	2.67	-5.39, 5.15	-0.05	.964
Child explains understanding of reading	-0.63	2.44	-5.46, 4.20	-0.26	.797
Child talks about what read	-0.10	2.67	-5.37, 5.16	0.04	.969

Table 20.

Summary Statistics of Fifth Grade Independent Variables used in the Second Research Question (n = 1880)

Variable	μ	*SD*	Range	Skewness	Kurtosis
Teacher age	42.61	5.95	26, 59	-0.08	2.67
Years with certified teacher	3.62	0.63	2, 4	-1.44	3.86
Elementary certification	2.81	0.36	2, 3	-1.91	4.64
Advanced degree	1.70	1.22	0, 4	0.27	2.09
Years teaching	14.83	6.66	1, 35	0.18	2.47
Choice					
Frequency child choose book	3.18	0.84	1, 4	-0.73	2.85
Child-selected activities	1.22	0.90	0, 3	0.27	2.30
Cross-curricular connections					
Integrate curriculum around themes	2.61	0.89	1, 4	-0.09	2.26
Read other subjects daily	0.34	0.48	0, 1	0.63	1.40
Discussion (frequency)					
Discuss interpretation of reading	0.18	0.38	0, 1	1.76	4.09
Discuss new vocabulary	2.78	0.88	1, 4	-0.23	2.32
Group projects about book	0.79	0.73	0, 2	0.31	1.92
Explain understanding of reading	0.33	0.45	0, 1	0.94	1.88
Child talks about what read	0.40	0.49	0, 1	0.39	1.15
Time on reading instruction	1.46	0.92	0, 3	0.14	2.19
Preparation to teach	0.94	0.21	0, 1	-4.40	20.34
Preparation to help	0.85	0.34	0, 1	-2.08	5.26

Note: These data were calculated by taking the mean of each variable across imputed datasets (M = 20).

Table 21.

Individual Contribution of Teacher Independent Variables used in the Second Research Question in Bivariate Regressions with Fifth Grade Reading IRT Scores (n = 1880)

Variables	*β*	*SE*	*95% CI*	*p* value
Teacher age	0.12	0.17	-0.23, 0.46	.505
Teacher race	**3.29**	**0.92**	**1.47, 5.10**	**<.001**
Years teaching	0.21	0.16	-0.10, 0.52	.191
Advanced degree	**2.13**	**0.98**	**0.19, 4.06**	**.032**
Choice			***F* (2, 244.9) = 3.47**	**.033**
Frequency child chooses books to read	**3.18**	**1.40**	**0.42, 5.95**	**.024**
Time on child selected activities	-2.13	1.22	-4.54, 0.27	.082
Cross-curricular connections			*F* (2, 211.0) = 1.07	.346
Integrate curricula around themes	0.32	1.24	-2.13, 2.77	.676
Read other subjects daily	3.01	2.94	-1.23, 7.10	.149
Discussion (frequency)			***F* (4, 294.6) = 2.56**	**.039**
Discuss new vocabulary	-0.93	1.28	-3.46, 1.60	.467
Group projects about a book	-0.14	1.46	-3.01, 2.73	.923
Child explains understanding of reading	5.98	2.06	1.91, 10.04	.004
Child talks about what read	**-5.35**	**2.25**	**-9.79, -0.90**	**.019**
Time on reading instruction	0.33	1.15	-1.95, 2.60	.777

Table 22.

Hierarchical Multiple Regression Analyses Predicting Fifth Grade Reading IRT Scores from Child and Teacher Independent Variables (n = 1880)

Independent Variables	ΔR	β	*SE*	*95% CI*	*F (df)*	*p value*
Step 1	0.36				**59.93 (6, 324.8)**	**< .001**
Teacher race		1.46	0.84	-0.18, 3.11		.081
Advanced degree		0.74	0.66	-0.57, 2.05		.266
Step 2	0.01				**57.62 (6, 317.5)**	**< .001**
Choice					1.88 (2, 220.1)	.154
Frequency child chooses books		1.39	0.97	-0.53, 3.31		.155
Time on child-selected activities		-1.17	0.94	-3.03, 0.69		.214
Step 3	0.01				**44.77 (8, 323.0)**	**< .001**
Discussion (frequency)					1.72 (4, 292.4)	.146
Discuss new vocabulary		-0.72	0.92	-2.53, 1.09		.435
Group projects about a book		-0.22	1.17	-2.52, 2.08		.853
Child explains understanding		3.48	1.93	-0.33, 7.30		.073
Child talks about what read		**-4.02**	**1.78**	**-7.52, -0.51**		**.025**

Table 23.

Descriptive Statistics of Kindergarten Within-Year Independent Variables

Variable	μ	*SD*	Range	Skewness	Kurtosis
Behavior engagement	2.82	0.79	1, 4	-0.02	2.22
Choice	1.20	0.75	0, 2	-0.35	1.84
Cross-curricular connections	0.97	0.18	0, 1	-5.13	27.30
Discussion	1.07	0.53	0, 2	0.07	3.48
Behavior/ choice	3.39	2.39	0, 8	0.20	2.21
Behavior/ cross-curricular	2.73	0.93	0, 4	-0.66	3.67
Behavior/ discussion	3.01	1.76	0, 8	0.64	4.12

Note. These data were calculated by taking the mean of each variable across imputed datasets (M = 20).

Table 24.

Individual Contribution of Kindergarten Independent Variables in Within-Year Bivariate Regressions with Kindergarten Reading IRT Scores as the Dependent Variable (n = 1880)

Variables	Bivariate Coefficients and Significance Statistics				
	β	*SE*	*t* value	95% *CI*	*p* value
Behavioral engagement	**1.40**	**0.27**	**5.23**	**0.87, 1.93**	**< .001**
Choice	0.05	0.24	0.21	-0.43, 0.53	.836
Discussion	**0.88**	**0.27**	**3.26**	**0.35, 1.41**	**.001**
Behavior/ discussion	**0.48**	**0.10**	**4.97**	**0.29, 0.67**	**< .001**
Behavior/ choice	**0.22**	**0.08**	**2.79**	**0.67, 0.39**	**.006**
Behavior/ cross-curricular	**1.18**	**0.21**	**5.51**	**0.76, 1.60**	**< .001**

Note. Final multiple regression model is displayed in Table 25.

Table 25.

Final Model of Multiple Regression Analyses Predicting Kindergarten Reading IRT Scores from Within-Year Behavioral Engagement, Instruction, and Interaction Variables (n = 1880)

	IRT Reading Scores				
Variables	*β*	*SE*	*t* value	95% *CI*	*p* value
Behavioral engagement	**2.37**	**0.63**	**3.79**	**1.14, 3.60**	**<.001**
Discussion	**3.26**	**1.28**	**2.55**	**0.74, 5.77**	**.011**
Behavior/ choice	0.01	0.32	0.03	-0.63, 0.65	.975
Behavior/ cross-curricular	0.25	0.27	0.95	-0.27, 0.78	.344
Behavior/ discussion	-0.86	0.45	-1.94	-1.74, 0.02	.054
Total Model		***F* (6, 298.3) =**	**9.16**		**< .001**

Table 26.

Descriptive Statistics of First Grade Within-Year Independent Variables (n = 1880)

Variables	μ	*SD*	Range	Skewness	Kurtosis
Behavioral engagement	2.77	0.82	1, 4	-0.06	2.30
Choice	1.26	0.60	0, 2	-0.18	2.44
Cross-curricular	0.66	0.48	0, 1	-0.66	1.43
Discussion	1.04	0.38	0, 2	0.43	6.59
Behavior/ choice	3.48	2.01	0, 8	0.46	2.95
Behavior/ cross-curricular	1.83	1.48	0, 4	-0.08	0.43
Behavior/ discussion	2.89	1.40	0, 8	0.86	5.72

Note. These data were calculated by taking the mean of each variable across imputed datasets (M = 20).

Table 27.

Individual Contribution of First Grade Independent Variables in Within-Year Bivariate Regressions with First Grade Reading IRT Scores as the Dependent Variable

	Bivariate Coefficients and Significance Statistics				
Variables	*β*	*SE*	*t* score	95% *CI*	*p* value
Behavioral engagement	**5.13**	**1.11**	**4.63**	**2.94, 7.33**	**< .001**
Choice	-0.15	1.34	-0.11	-2.81, 2.50	.910
Cross-curricular connections	-0.88	1.65	-0.54	-4.15, 2.37	.592
Discussion	0.23	2.04	0.11	-3.82, 4.29	.910
Behavior/ discussion	**1.90**	**0.55**	**3.48**	**0.82, 2.98**	**.001**
Behavior/ choice	**1.06**	**0.42**	**2.55**	**-0.24, 1.89**	**.012**
Behavior/ cross-curricular	0.80	0.56	1.42	-0.31, 1.91	.156

Note. Final multiple regression model is displayed in Table 28.

Table 28.

Final Model of Multiple Regression Analyses Predicting First Grade Reading IRT Scores from Within-Year Behavioral Engagement and Interaction Variables (n = 1880)

	IRT Reading Scores			
Independent Variables	*β*	*SE*	*t* score	*p* value
Behavioral engagement	**4.39**	**1.38**	**3.17**	**.002**
Behavior/ choice	-0.38	0.51	-0.08	.940
Behavior/ discussion	0.13	0.67	0.20	.840
Total Model		***F* (3, 166.1) = 6.50**		**< .001**

Table 29.

Descriptive Statistics of Third Grade Within-Year Independent Variables (n = 1880)

Variables	μ	*SD*	Range	Skewness	Kurtosis
Behavioral engagement	2.73	0.83	1, 4	-0.04	2.29
Choice	1.09	0.56	0, 2	0.03	3.08
Cross-curricular connections	1.02	0.75	0, 2	-0.05	1.76
Discussion	0.79	0.49	0, 2	-0.38	3.10
Behavior/ choice	3.00	1.85	0, 8	0.63	3.64
Behavior/ cross-curricular	2.82	2.32	0, 8	0.48	2.48
Behavior/ discussion	2.17	1.57	0, 8	0.27	3.39

Note. These data were calculated by taking the mean of each variable across imputed datasets (M = 20).

Table 30.

Individual Contribution of Third Grade Independent Variables in Within-Year Bivariate Regressions with Third Grade Reading IRT Scores as the Dependent Variable (n = 1880)

	Bivariate Coefficients and Significance Statistics				
Variables	*β*	*SE*	*t* score	95% *CI*	*p* value
Behavioral engagement	**6.02**	**1.48**	**4.06**	**3.08, 8.96**	**< .001**
Choice	-1.50	2.27	-0.66	-5.98, 2.98	.510
Cross-curricular connections	-1.14	1.59	-0.71	-4.26, 1.99	.476
Discussion	-2.74	2.33	-1.18	-7.33, 1.85	.240
Behavior/ discussion	0.67	0.76	0.88	-0.83, 2.16	.379
Behavior/ choice	0.89	0.62	1.45	-0.33, 2.13	.149
Behavior/ cross-curricular	0.20	0.45	0.45	-0.69, 1.09	.655

Table 31.

Descriptive Statistics of Fifth Grade Within-Year Independent Variables (n = 1880)

Variable	μ	*SD*	Range	Skewness	Kurtosis
Behavioral engagement	2.77	0.76	1, 4	0.08	2.30
Choice	0.83	0.55	0, 2	-0.06	2.94
Cross-curricular connections	0.80	0.77	0, 2	0.35	1.77
Discussion	0.67	0.52	0, 2	-0.22	2.13
Behavior/ choice	2.32	1.72	0, 8	0.42	3.53
Behavior/ cross-curricular	2.23	2.29	0, 8	0.74	2.70
Behavior/ discussion	1.88	1.58	0, 8	0.24	2.63

Table 32.

Individual Contribution of Fifth Grade Independent Variables in Within-Year Bivariate Regressions with Fifth Grade Reading IRT Scores as the Dependent Variable (n = 1880)

	Bivariate Coefficients and Significance Statistics				
Variables	*β*	*SE*	*t* score	95% *CI*	*p* value
Behavioral engagement	**7.21**	**1.58**	**4.57**	**4.09, 10.33**	**< .001**
Choice	3.79	2.17	1.74	-0.49, 8.07	.082
Cross-curricular connections	1.93	1.22	1.58	-0.48, 4.34	.116
Discussion	0.60	1.84	0.33	-3.02, 4.22	.745
Behavior/ discussion	1.21	0.61	1.97	-0.003, 2.42	.051
Behavior/ choice	**2.10**	**0.77**	**2.72**	**0.58, 3.62**	**.007**
Behavior/ cross-curricular	**1.20**	**0.40**	**3.03**	**0.42, 1.98**	**.003**

Note. Final multiple regression model is displayed in Table 33.

Table 33.

Final Model of Multiple Regression Analyses Predicting Fifth Grade Reading IRT Scores from Within-Year Behavioral Engagement, Instruction, and Interaction Variables (n = 1880)

	IRT Reading Scores				
Independent Variables	*β*	*SE*	*t* score	95% *CI*	*p* value
Behavioral engagement	**5.97**	**1.66**	**3.60**	**2.69, 9.26**	**< .001**
Behavior/ choice	0.70	0.79	0.89	-0.85, 2.25	.375
Behavior/ cross-curricular	0.52	0.43	1.22	-0.32, 1.36	.223
Total Model		***F* (3, 255.5) = 9.26**			**< .001**

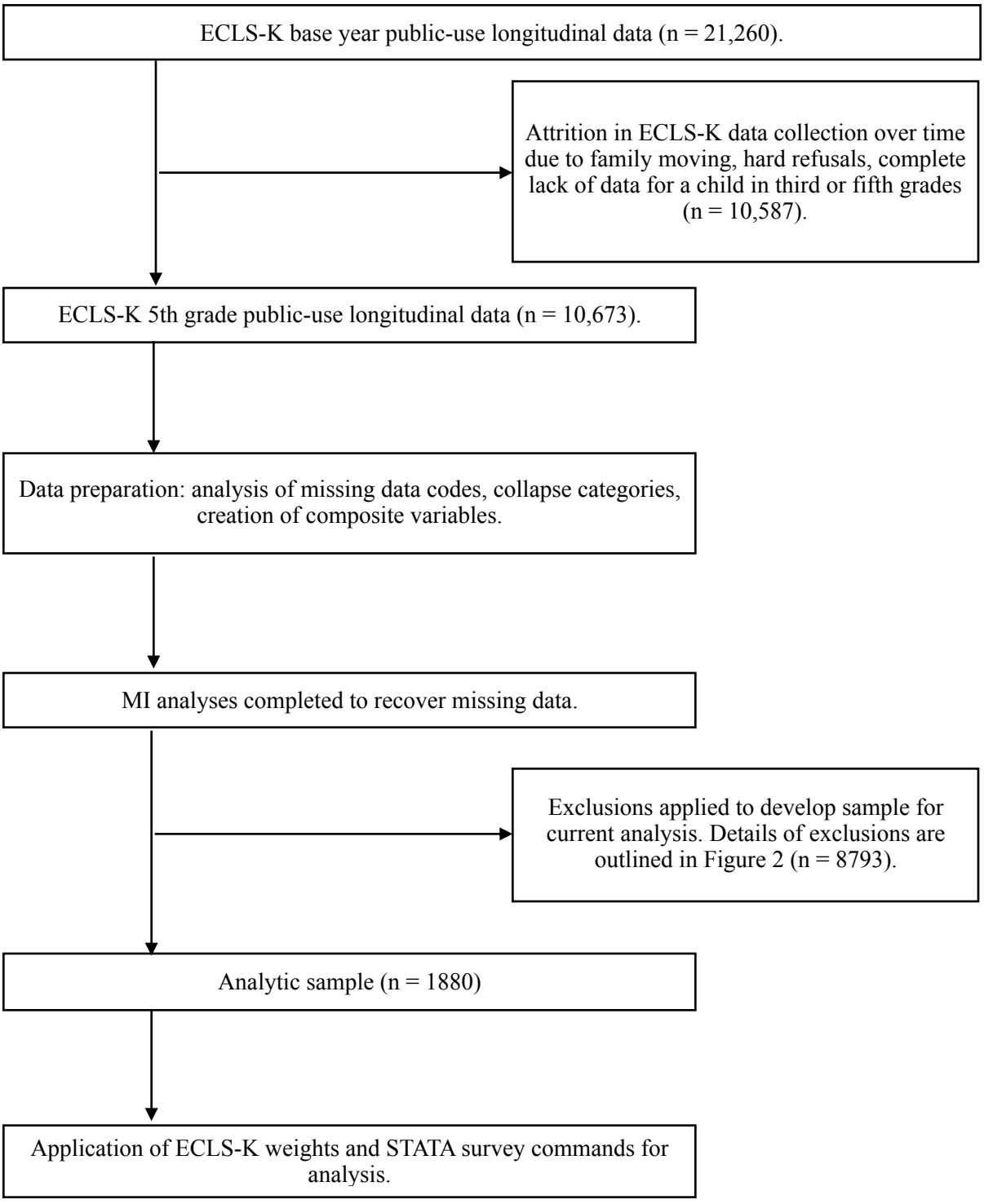
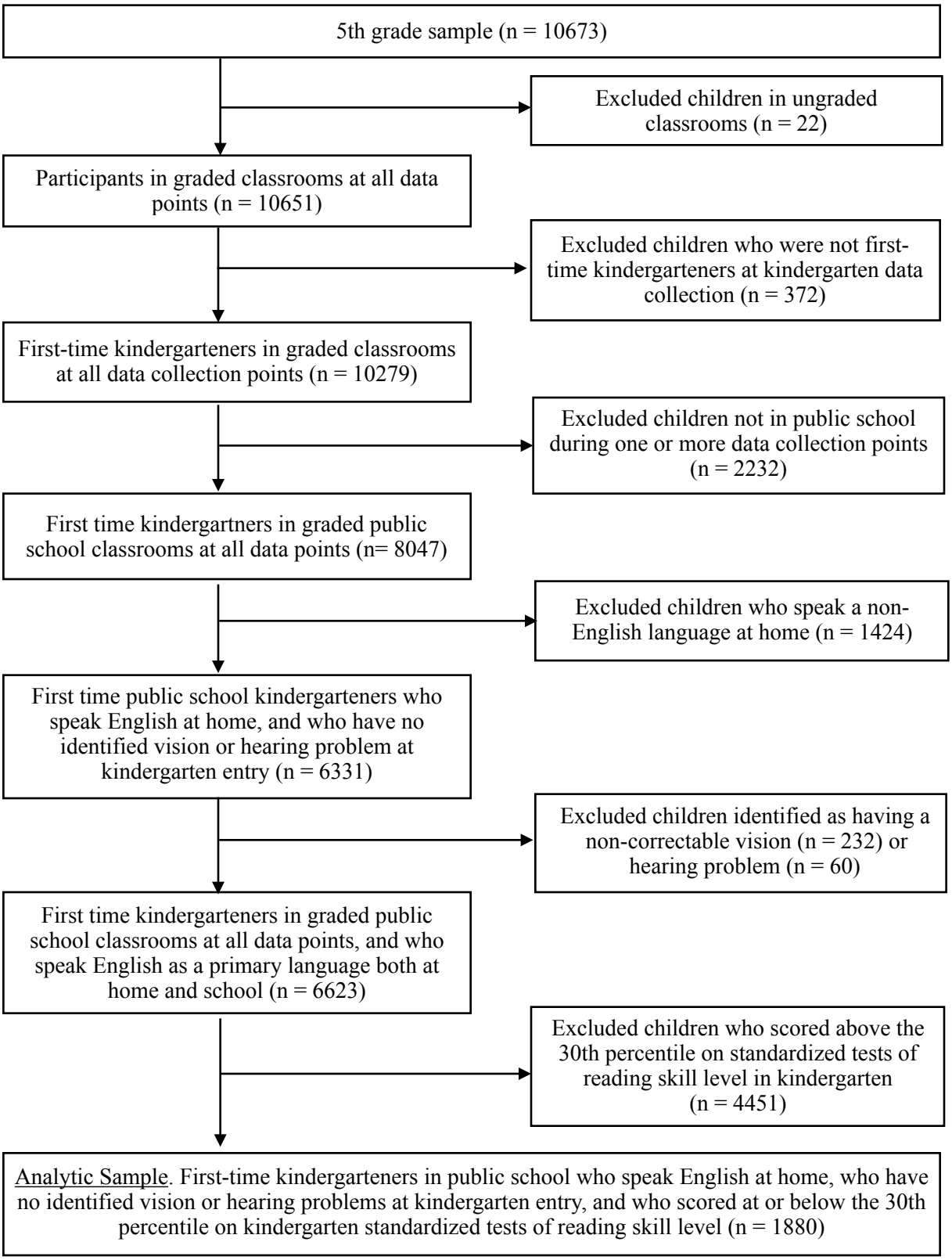


Figure 1. Flow chart of the development of the analytic sample from the fifth grade public-use longitudinal dataset.



F*igure 2:* Flow chart of sample exclusions.

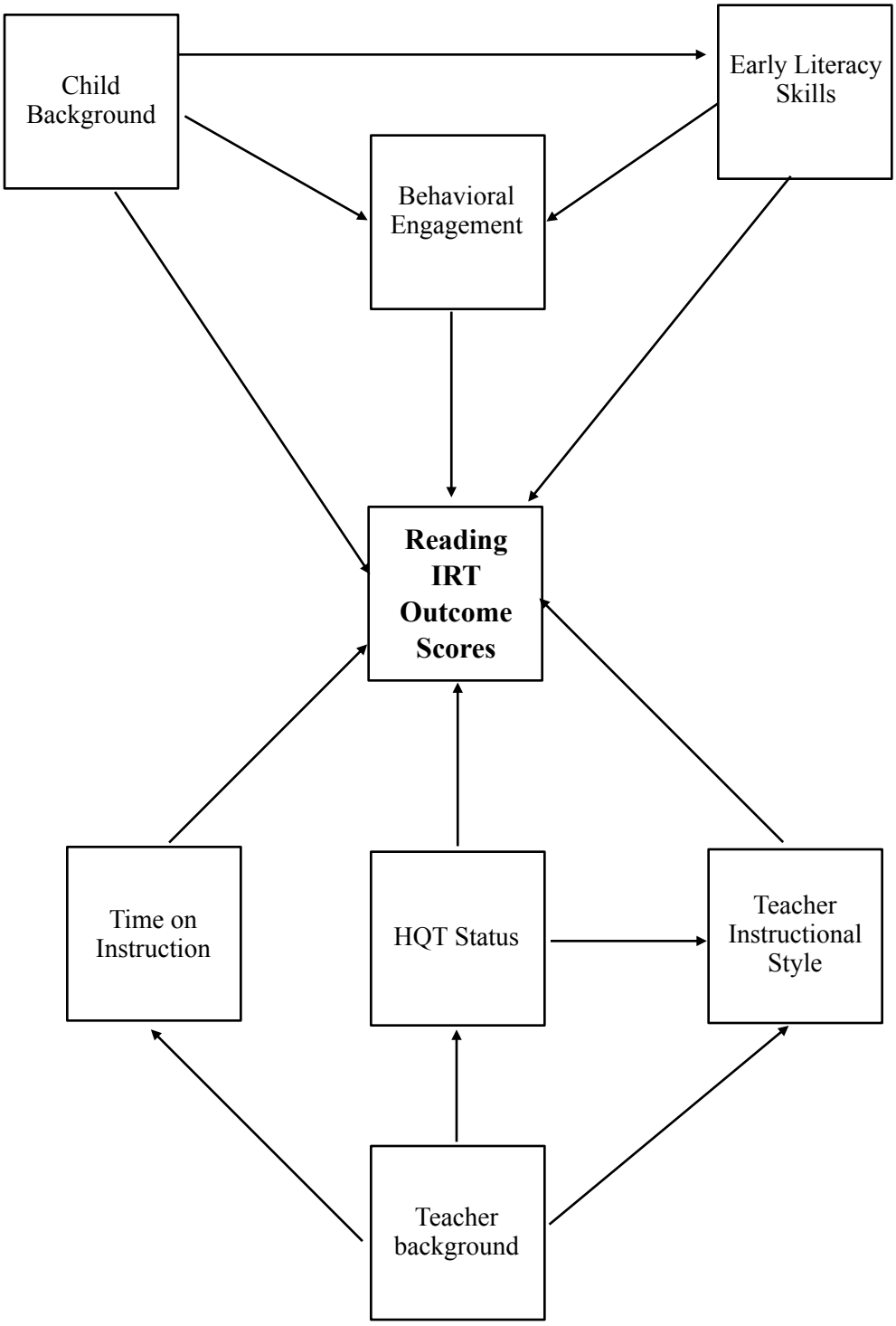


Figure 3: Conceptual model of research study.

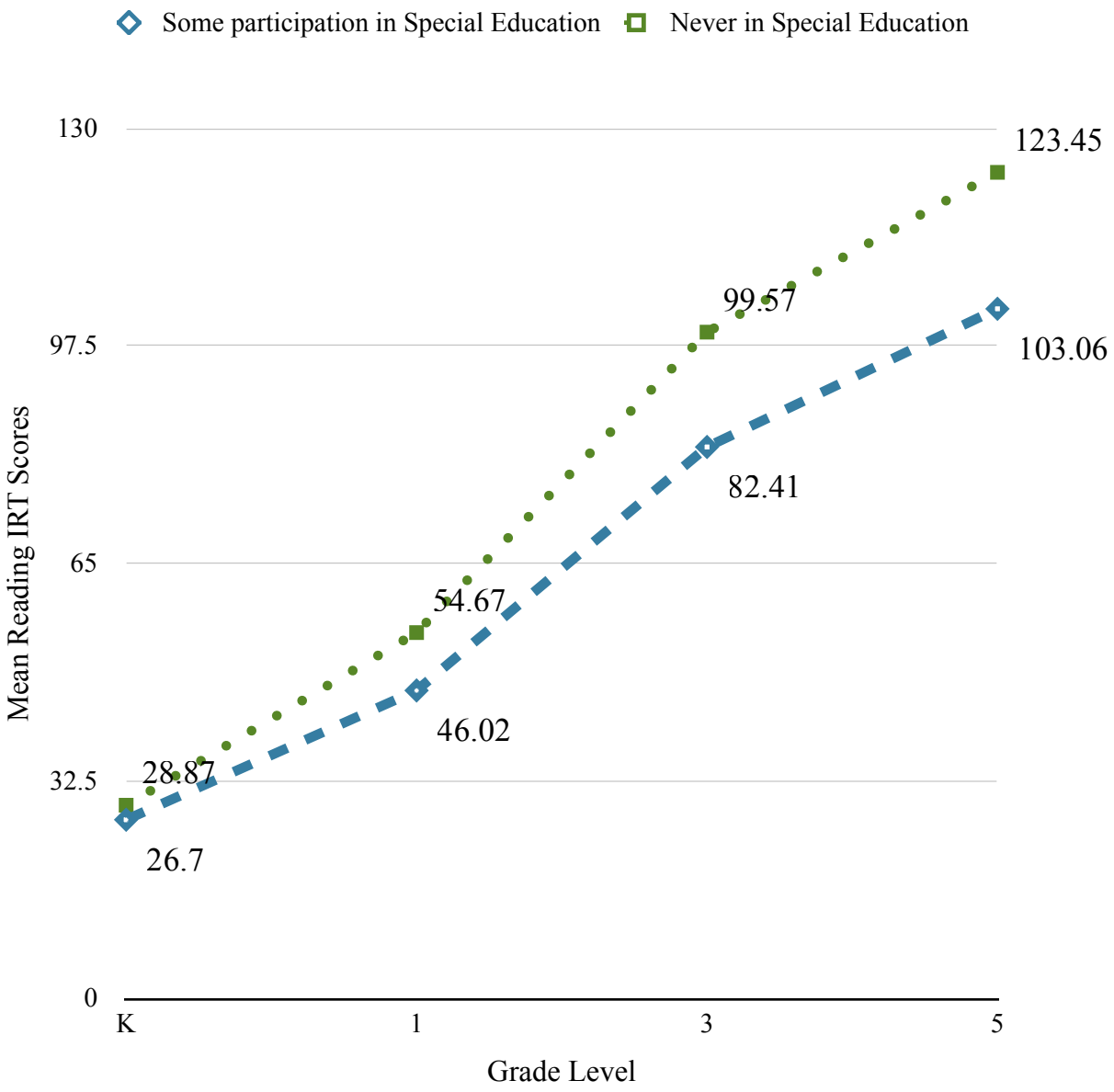


Figure 4: Sampled participants' mean IRT scores across years of data collection and status of participation in special education services.

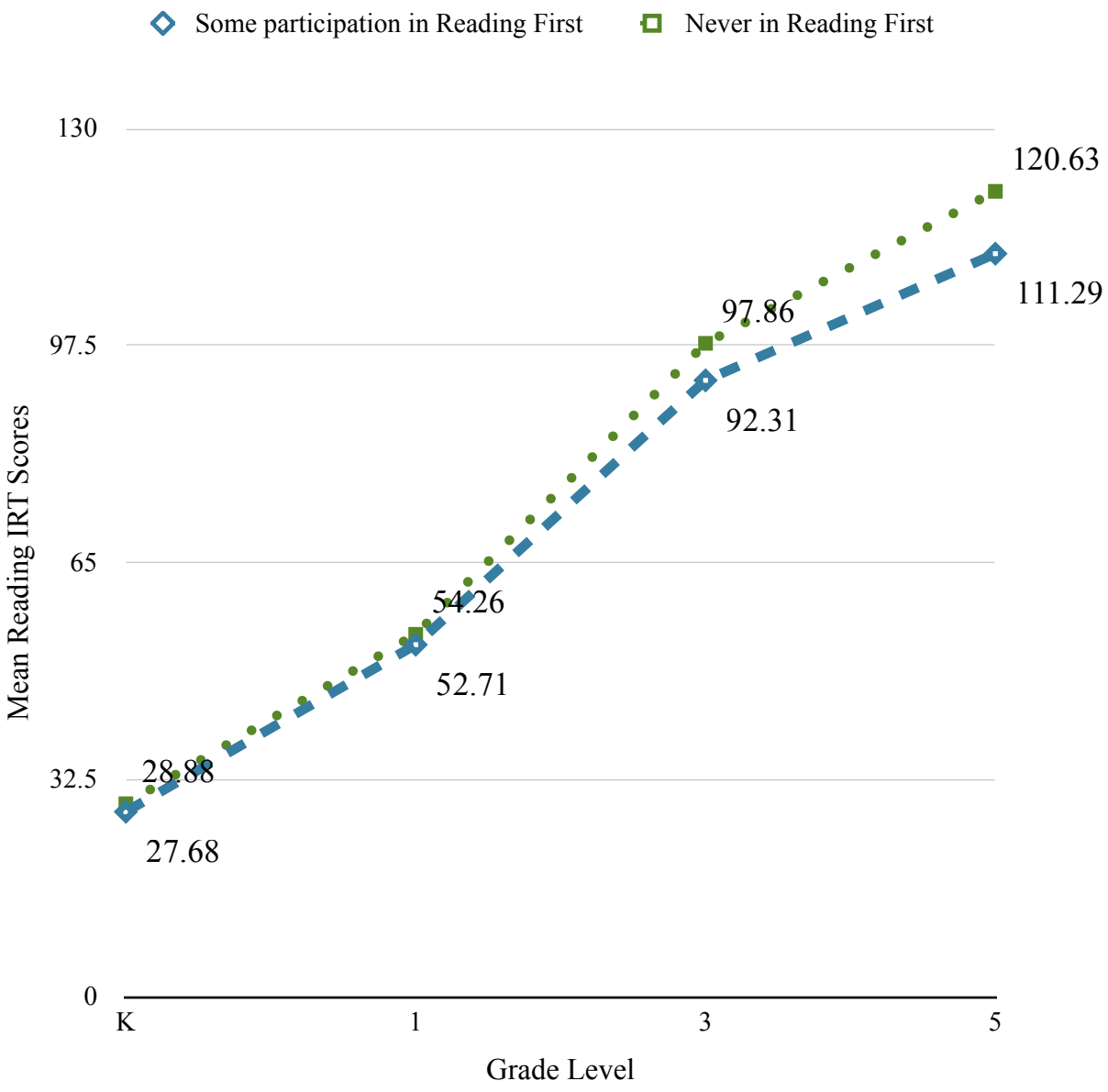


Figure 5: Sampled participants' mean IRT scores across years of data collection and status of participation in Reading First services.

REFERENCES

Agostin, T. M., & Bain, S. K. (1997). Predicting early school success with developmental and social skills screeners. *Psychology in the Schools, 34,* 219-228. doi: 10.1002(SICI)1520-6807(199707)34:3<219::AID-PITS4>3.0.CO;2-

Akbari, R., & Dadvand, B. (2011). Does Formal Teacher Education Make a Difference? A Comparison of Pedagogical Thought Units of B.A. versus M.A. Teachers. *The Modern Language Journal, 95*(1), 44-60. doi: 10.1111/j.1540-4781.2010.01142.x0026-7902/11/44-60

Alexander, K. L., & Entwisle, D. R. (1988). Achievement in the first 2 years of school: Patterns and processes. Chicago: University of Chicago Press.

Alexander, K. L., Entwisle, D. R., & Dauber , S. L. (1993). First-grade classroom behavior: Its short- and long-term consequences for school performance. *Child Development, 64,* 801-814. Retrieved from: http://www.jstor.org.libproxy.lib.unc.edu/stable/1131219

Allison, P. D. (1999). Multiple Regression: A Primer. Thousand oaks, CA: Sage Publications, Inc.

Allison, P. D. (2002). Missing Data. Thousand oaks, CA: Sage Publications, Inc. doi: 10.4135/9781412985079

Alvermann, D. E., Hagood, M. C., Heron, A., Hughes, P., Williams, K., & Yoon, J. (2002). *After-school media clubs for struggling adolescent readers: A study of youth's critical awareness* (Grant #199900278). Athens, GA: University of Georgia.

Alvermann, D. E., Hagood, M. C., Heron-Hruby, A., Hughes, P., Williams, K. B., Yoon, J.-C. (2007). Telling themselves who they are: What one out-of-school time study revealed about underachieving readers. Reading Psychology, 28, 31-50. doi: 10.1080/02702710601115455

Anderson, G. E., Whipple, A. D., & Jimerson, S. R. (2002). Grade retention: Achievement and mental health outcomes. In N. A. o. S. Psychologists (Ed.), *NASP Toolkit: Practical resources at your fingertips*. Bethesda, MD: Author.

Archambault, I., Pagani, L. S., & Fitzpatrick, C. (2013). Transactional associations between classroom engagement and relations with teachers from first through fourth grade. *Learning and Instruction, 23*, 1-9. doi: http://dx.doi.org/10.1016/j.learninstruc.2012.09.003

Baroody, A. E., Rimm-Kaufman, S. E., Larsen, R. A., Curby, T. W. (2014). The link between responsive classroom training and student-teacher relationship quality in the fifth grade: A study of fidelity of implementation. *School Psychology Review, 43*, 69-85.

Bast, J. & Reitsma, P. (1997). Matthew effects in reading: A comparison of latent growth curve models and simplex models with structured means. *Multivariate Behavioral Research, 32,* 135-167. Retrieved from:

Baydar, N., Brooks-Gunn, J., & Furstenberg, F. F., (1993). Early warning signs of functional illiteracy: Predictors in childhood and adolescence. *Child Development, 64*, 815-829. doi: 10.1111/1467-8624.ep9308115030

Benner, G. J., Beaudoin, K., Kinder, D., & Mooney, P. (2005). The relationship between the beginning reading skills and social adjustment of a general sample of elementary aged children. *Education and Treatment of Children, 28*, 250-264. Retrieved from:

Bestor, J. (2014). *Viewpoint: Common Core Standards do not serve the educational needs of our children.* NASP Communique, 42(6), 22.

Bishop, A. G. (2003). Prediction of first-grade reading achievement: A comparison of fall and winter kindergarten screening. *Learning Disability Quarterly, 26,* 189-200. Retrieved from: http://www.jstor.org/stable/1593651

Block, C. C., Parris, S. R., Reed, K. L., Whiteley, C. S., & Cleveland, M.D. (2009). Instructional approaches that significantly increase reading comprehension. *Journal of Educational Psychology, 101*, 262-281. doi: 10.1037/a0014319

Brock, I. L., Nishida, T. K., Chiong, C., Grimm, K. J., & Rimm-Kaufman, S. E. (2008). Children's perceptions of the classroom environment and social and academic performance: A longitudinal analysis of the contribution of the Responsive Classroom approach. *Journal of School Psychology, 46*, 129-149.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Psychology* (pp. 993-1028). New York, NY: John Wiley & Sons, Inc.

Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328-375). New York: Macmillan.

Butler, S. R., Marsh, H. W., Sheppard, M. J., & Sheppard, J. L. (1985). Seven-year longitudinal study of the early prediction of reading achievement. *Journal of Educational Psychology, 77,* 349-361.

Cairns, R. B., Cairns, B. D., & Neckerman, H. J. (1989). Early school dropout: Configurations and determinants. *Child Development, 60,* 1437-1452. doi: 0009-3920/89/6006-0018

Campbell, F. A., Ramey, C. T., Pungello, E. P., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian project. *Applied Developmental Science, 6,* 42-57.

Carlisle, J. F., Correnti, R., Phelps, G., & Zeng, J. (2009). Exploration of the contribution of teachers' knowledge about reading to their students' improvement in reading. *Reading & Writing, 22*(4), 457-486. doi: 10.1007/s11145-009-9165-y

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the Early Childhood Longitudinal Study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology, 98*, 489-507. doi: 10.1037/0022-0663.98.3.489

Christ, T. Nelson, P. & Demers, J. (2013). Responsive Environmental Assessment for Classroom Teaching. Retrieved on March 30, 2014 from: www.mspaonline.net/.../Thurs-REACT-2014-Christ-Nelson-Demers1.doc

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review, 28*, 415-427. doi: 10.1016/j.econedurev.2008.09.003

Clark, K.F., & Graves, M.F. (2005). Scaffolding students' comprehension of text. *The Reading Teacher, 58*, 570-580.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/ correlational analysis for the behavioral sciences* (Third ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Coie, J. D., Watt, N. F., West, S. G., Hawkins, J. D., Asarnow, J. R., Markman, H. J., … Long, B. (1993). The science of prevention: A conceptual framework and some directions for a national research program. *American Psychologist, 48*, 1013-1022.

Connor, C. M., Morrison, F. J., Fishman, B., Giuliani, S., Luck, M., Underwood, A. B., . . . Schatschneider, C. (2011). Testing the Impact of Child Characteristics x Instruction Interactions on Third Graders' Reading Comprehension by Differentiating Literacy Instruction. *Reading Research Quarterly, 46*(3), 189-221.

Connor, C. M., Son, S.-H., Hindman, A. H., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology, 43,* 343-375. doi: 10.1016/j.jsp.2005.06.001

Croninger, R. G., Rice, J. K., Rathbun, A., & Nisho, M. (2003). *Teacher qualifications and first grade achievement: A multilevel analysis* (CEPAL Occasional Paper OP-03-2). Retrieved from University of Maryland, Center for Education, Policy and Leadership website: *http://www.*education.umd.edu/EDPA/CEPAL/pdf/oc-papers/CEPAL-OP-03-2.pdf

Cunningham, A.E., & Stanovich, K.E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33,* 934-945.

Datar, A. (2006). Does delaying kindergarten entrance give children a head start? *Economics of Education Review, 25*, 43-62.

Demaray, M. K., Ruffalo, S. L., Carlson, J., Busse, R. T., Olson, A. E., McManus, S. M., & Leventhal, A. (1995). Social skills assessment: A comparative evaluation of six published rating scales. *School Psychology Review, 24*(4), 648-671.

Denton, K., & West, J. (2002). *Children's reading and mathematics achievement in kindergarten and first grade* (NCES Publication No. 2002-125). Washington, DC: National Center for Education Statistics, U.S. Department of Education, Office of Educational Research and Improvement.

Denton, K., West, J., & Walston, J. (2003). *Reading- Young children's achievement and classroom experiences.* (NCES Publication No. 2003-070). Washington, DC: National Center for Education Statistics, U.S. Department of Education, Office of Educational Research and Improvement.

Desimone, L. M. (2010). Teacher effects and the achievement gap: Do teacher and teaching quality influence the achievement gap between black and white and high- and low-SES students in the early grades? *Teachers College Record (1970), 112*(12), 3024.

DiPerna, J. C., Lei, P.-W., & Reid, E. E. (2007). Kindergarten predictors of mathematical growth in the primary grades: An investigation using the early childhood longitudinal study – kindergarten cohort. *Journal of Educational Psychology, 99,* 369-379. doi: 10.1037/0022-0663.99.2.369

Doumen, S., Koomen, H. M., Buyse, E., Wouters, S., & Verschueren, K. (2012). Teacher and observer views on student-teacher relationships: Convergence across kindergarten and relations with student engagement. *Journal of School Psychology, 50*, 61-76.

Downer, J. T., Rimm-Kaufman, S. E., & Pianta, R. C. (2007). How do classroom conditions and children's risk for school problems contribute to children's behavioral engagement in learning? *School Psychology Review, 36,* 413-432.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., … Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43,* 1428-1446. doi: 10.1037/0012-1649.43.6.1428

Eamon, M. K. (2002). Effects of poverty on mathematics and reading achievement of young adolescents. *Journal of Early Adolescence, 22*, 49-74.

Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., … Barbarin, O. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly, 21,* 174-195. doi: 10.1016/j.ecresq.2006.04.004

Early, D. M., Maxwell, K. L., Burchinal, M. R., Alva, S., Bender, R. H., Bryant, D., … Vandergrift, N. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development, 78*, 558-580.

Eddings, W. (2011, February 3, 2011). [Personal email request.].

Elbro, C., & Peterson, D. K. (2004). Long-term effects of phoneme awareness and letter sound training: An intervention study with children at risk for dyslexia. *Journal of Educational Psychology, 96,* 660-670. doi: 10.1037/0022-0663.96.4.660

Elder, T. E. & Lubotsky, D. H. (2009). Kindergarten entrance age and children's achievement. *Journal of Human Resources, 44*, 641-683.

Elliott, S. (1993). *Caring to learn: A report on the positive impact of a social curriculum.* Greenfield, MA: Northeast Foundation for Children. Retrieved from http://www.responsiveclassroom.org/pdf_files/caring_to_learn.pdf

Entwisle, D. R. & Alexander, K. L. (1993). Entry into school: The beginning school transition and educational stratification in the United States. *Annual Review of Sociology, 19*, 401-423.

Fantuzzo, J., Bulotsky-Shearer, R., McDermott, P. A., McWayne, C., Frye, D., & Perlman, S. (2007). Investigation of dimensions of social-emotional classroom behavior and school readiness for low-income urban preschool children. *School Psychology Review, 36,* 44-62.

Food and Nutrition Service. (2013). *Program Fact Sheet: National School Lunch Program.* Alexandria, VA: U.S. Department of Agriculture. Retrieved January 30, 2014, from http://www.fns.usda.gov/nslp/national-school-lunch-program

Fraser, B.J. (1998). Classroom environment instrucments: Development, validity, and applications. *Learning Environments Research, 1*, 7-33.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Review of Educational Research, 74,* 59-109.

Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., Unlu, F., Bozzi, L., … Rosenblum, S. (2008). *Reading First impact study: Final report.* (NCEE Publication No. 2009-4038). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute for Education Sciences, U.S. Department of Education.

Gilbert, J. K., Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Barquero, L. A., & Cho, E. (2013). Efficacy of a First-Grade Responsiveness-to-Intervention Prevention Model for Struggling Readers. *Reading Research Quarterly, 48*(2), 135-154.

Goldhaber, D. D. & Brewer, D. J. (1997). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *The Journal of Human Resources, 32(3)*. Retrieved from http://ssrn.com/abstract=3673.

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(2),* 129-145.

Greenwood, C. R., Horton, B. T., & Utley, C. A. (2002). Academic engagement: Current perspectives on research and practice. *School Psychology Review, 31,* 328-349.

Gresham, F. M., & Elliott, S. N. (1990a). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.

Gresham, F. M., & Elliott, S. N. (1990b). *Elementary Scale A ("How Often?)*. Circle Pines, MN: American Guidance Service.

Guarino, C. M., Hamilton, L. S., Lockwood, J. R., Rathbun, A. H., Germino-Hausken, E. (2006). *Teacher qualifications, instructional practices and reading and mathematics gains of kindergarteners* (NCES Publication No. 2006-031). Washington, DC: National Center for Education Statistics, U.S. Department of Education, Office of Educational Research and Improvement.

Guthrie, J. T., Klauda, S. L., & Ho, A. N. (2013). Modeling the Relationships Among Reading Instruction, Motivation, Engagement, and Achievement for Adolescents. *Reading Research Quarterly, 48*(1), 9-26.

Hall, L.A. (2009). Struggling reader, struggling teacher: An examination of student-teacher transactions with reading instruction and text in social studies. *Research in the Teaching of English, 43*, 286-309.

Hall, L. A. (2010). The negative consequences of becoming a good reader: Identity theory as a lens for understanding struggling readers, teachers, and reading instruction. *Teachers College Record, 112*, 1792-1829.

Hall, L.A., Johnson, A., Juzwik, M.M., Wortham, S., & Mosley, M. (2010). Teacher identity in the context of literacy teaching: Three explorations of classroom positioning and interaction in secondary schools. *Teaching and Teacher Education, 26*, 234-243.

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, *76*, 949-967.

Hanson, R. A., & Farrell, D. (1995). The long-term effects on high school seniors on learning to read in kindergarten. *Reading Research Quarterly, 30*, 908-933.

Hawk, P. P., Coble, C. R., & Swanson, M. (1985). Certification: It does matter. *Journal of Teacher Education, 36*. doi: 10.1177/002248718503600303

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*, 399-432. doi: 10.1177/0013161X07306452

Henk, W. A., & Melnick, S. A. (1995). The Reader Self-Perception Scale (RSPS): A new tool for measuring how children feel about themselves as readers. *Reading Teacher, 48*, 470-482.

Hong, G., & Yu, B. (2008). Effects of kindergarten retention on chidlren's social-emotional development: An application of propensity score method to multivariate, multilevel data. *Developmental Psychology, 44*, 407-421.

Houtveen, A. A. & van de Grift, W. J. (2007). Effects of metacognitive strategy instruction and instruction time on reading comprehension. *School Effectiveness and School Improvement, 18(2),* 173-190. doi: 10.1080/09243450601058717

Howse, R. B., Lange, G., Farran, D. C., & Boyles, C. D. (2003). Motivation and self-regulation as predictors of achievement in economically disadvantaged young children. *The Journal of Experimental Education, 71(2),* 151-174.

Hughes, J., & Kwok, O.-M. (2007). Influence of student-teacher and parent-teacher relationships on lower achieving readers' engagement and achievement in the primary grades. *Journal of Educational Psychology, 99*, 39-51. doi: 10.1037/0022-0663.99.1.39

Hughes, J. N., Kwok, O.-M., & Im, M. H. (2013). Effect of retention in first grade on parents' educational expectations and children's academic outcomes. *American Educational Research Journal, 50*(6), 1336-1359. doi: 10.3102/0002831213490784

Im, M. H., Hughes, J. N., Kwok, O.-M., Puckett, S., & Cerda, C. A. (2013). Effect of retention in elementary grades on transition to middle school. *Journal of School Psychology, 51*, 349-365.

Jimerson, Anderson, & Whipple. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools, 39*, 441-457.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437-447.

Justice, L. M., Invernizzi, M., Geller, K., Sullivan, A. K., & Welsch, J. (2005). Descriptive-developmental performance of at-risk preschoolers on early literacy tasks. *Reading Psychology, 26*, 1-25. doi: 10.1080/02702710490897509

Kaiser, J., Retelsdorf, J., Sudkamp, A., & Moller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction, 28*, 73-84.

Kennedy, E. (2010). Improving Literacy Achievement in a High-Poverty School: Empowering Classroom Teachers Through Professional Development. *Reading Research Quarterly, 45*(4), 384-387.

Kirby, J. R., Parrila, R. K., & Pfeiffer, S. L. (2003). Naming speed and phonological awareness as predictors of reading development. Journal of Educational Psychology, 95, 453-464. doi: 10.1037/0022-0663.95.3.453

Korhonen, J., Linnanmaki, K., & Aunio, P. (2014). Learning difficulties, academic well-being and educational dropout: A person-centered approach. *Learning and Individual Differences, 31,* 1-10. doi: http://dx.doi.org/10.1016/j.lindif.2013.12.011

Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Economics of Education Review, 28,* 49-57. doi: 10.1016/j.econedurev.2007.10.007

Lane, K. L., O'Shaughnessy, T. E., Lambros, K. M., Gresham, F. M., & Beebe-Frankenberger, M. E. (2001). The efficacy of phonological awareness training with first-grade students who have behavior problems and reading difficulties. *Journal of Emotional and Behavioral Disorders, 9,* 219-231.

Lee, E. S. & Forthofer, R. N. (2006). Strategies for variance estimation, pp. 22-39. Analyzing Complex Survey Data, 2nd ed. Sage Publications.

Lewis, J. (1979). Teacher training, teacher experience and pupil achievement. *Journal of instructional psychology, 6*(4), 3.

Lincove, J. A. & Painter, G. (2006). Does the age that children start kindergarten matter? Evidence of long-term educational and social outcomes. *Educational Evaluation and Policy Analysis, 28*, 153-179.

Livingston, A., & Wirt, J. (2004). *The condition of education in brief* (NCES Publication No. 2004-076). Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology, 36,* 596-613. doi: 10.1037/0012-1649.36.5.596

Malecki, C. K., & Elliott, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly, 17,* 1-23.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., … Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development, 79*, 732-749.

Mather, N., & Ofiesh, N. (2005). Resilience and the child with learning disabilities. In S. Goldstein and R. Brooks (Eds.), *Handbook of resilience in children* (pp. 239-255). New York, NY: Kluwer Academic/ Plenum Publishers.

Meisels, S. J. (1992). Doing harm by doing good: Iatrogenic effects of early childhood enrollment and promotion policies. *Early Childhood Research Quarterly, 7*, 155 - 174.

Melekoglu, M.A. (2011). Impact of motivation to read on reading gains for struggling readers with and without learning disabilities. *Learning Disability Quarterly, 34*, 248-261.

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly, 21,* 471-490. doi: 10.1016/j.ecresq.2006.09.003

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43,* 947-959. doi: 10.1037/0012-1649.43.4.947

McClelland, M. M., Morrison, F. J., & Holmes, D. L. (2000). Children at risk for early academic problems: The role of learning-related social skills. *Early Childhood Research Quarterly, 15,* 307-329.

McCoach, D. B., O'Connell, A. A., Reis, S. M., & Levitt, H. A. (2006). Growing readers: A hierarchal linear model of children's reading growth during the first two years of school. *Journal of Educational Psychology, 98,* 14-28. doi: 10.1037/0022-0663.98.1.14

McWayne, C. M., Fantuzzo, J. W., & McDermott, P. A. (2004). Preschool competency in context: An investigation of the unique contribution of child competencies to early academic success. *Developmental Psychology, 40*, 633-645. doi: 10.1037/0012-1649.40.4.633

Moos, R. H. & Trickett, E.J. (1987). Classroom Environment Scale manual (2nd ed.) Palo Alto, CA: Consulting Psychologists Press.

Moser, S. E., West, S. G., & Hughes, J. N. (2012). Trajectories of math and reading achievement in low-achieving children in elementary school: Effects of early and later retention in grade. *Journal of Educational Psychology, 104*(3), 603-621. doi: 10.1037.a0027571

Munoz, M. M., & Chang, F. C. (2007). The elusive relationship between teacher characteristics and student academic growth: A longitudinal multilevel model for change. *Journal of Personnel Evaluation in Education, 20*, 147-164. doi: 10.1007/s11092-008-9054-y

NAEYC (1998). A joint position statement by NAEYC & International Reading Assoc.: Learning to read and write: Developmentally appropriate practices for young children. *Young Children, 53*(4), 30-46.

NASP (2011). Position Statement: *Grade Retention and Social Promotion*. NASP: Bethesda, MD.

NCES (2006). *ECLS-K Longitudinal Kindergarten-Fifth Grade Public-Use Child File: Data Files and Electronic Code Book*. NCES2006-035 Washington DC: U. S. Department of Education, Institute of Educational Sciences.

NCES (2014). *The nation's report card: A First Look: 2013 Mathematics and Reading* (NCES 2014-451). Institute for Education Sciences, U.S. Department of Education: Washington, DC.

National Governors Association Center for Best Practices & Council of Chief State School Officers (NGACBP). (2010). Common Core State Standards. Washington, DC: Authors.

NICHD (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, D.C.: U.S. Government Printing Office.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

Northeast Foundation for Children (2007). *Guidelines for the responsive classroom.* Greenfield, MA: Author.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*, 237-257.

O'Connor, R. E., & Jenkins, J. R. (1999). Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading, 3(2),* 159-197.

Otaiba, S. A., Connor, C., Lane, H., Kosanovich, M. L., Schatschneider, C., Dyrlund, A. K., … Wright, T. L. (2008). Reading First kindergarten classroom instruction and students' growth in phonological awareness and letter naming-decoding fluency. *Journal of School Psychology, 46,* 281-314. doi:10.1016/j.jsp.2007.06.002

Parlardy, G. J., & Rumberger, R. W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis, 30(2),* 111-140. doi:10.3102/0162373708317680

Phillips, L. M., Norris, S. P., Osmond, W. C., & Maynard, A. M. (2002). Relative reading achievement: A longitudinal study of 187 children from first through sixth grades. *Journal of Educational Psychology, 94,* 3-13. doi: 10.1037//0022-0663.94.1.3

Pianta, R. C., La Paro, K., & Hamre, B. (2007). *Classroom Assessment Scoring System* (CLASS) *manual*, K-3. Baltimore, MD: Brookes.

Pollack, J. M., Atkins-Burnett, S., Najarian, M., & Rock, D. A. (2005). *ECLS-K Psychometric report for the fifth grade* (NCES 2006-036). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Ponitz, C. C., Rimm-Kaufman, S. E., Grimm, K. J., & Curby, T. W. (2009). Kindergarten classroom quality, behavioral engagement, and reading achievement. *School Psychology Review, 38,*102-120.

Pressley, M., Wharton-McDonald, R., Allington, R., Block, C. C., Morrow, L., Tracey, D., … Woo, D. (2001). A study of effective first-grade literacy instruction. *Scientific Studies of Reading, 5*, 35-58.

Rao, J. N. (1996). On variance estimation with imputed survey data. *Journal of the American Statistical Association, 91(434), 499-506.* doi: 10.1080/01621459.1996.10476910

Rathbun, A., West, J., and Walston, J. (2005, April). *Relationships between family risks and children's reading and mathematics growth from kindergarten through third grade.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Reutzel, D.R., Petscher, Y., & Spichtig, A.N. (2012). Exploring the value added of a guided silent reading intervention: Effects on struggling third grade readers' achievement. *The Journal of Educational Research, 105*, 404-415.

Rimm-Kaufman, S.E., Fan, X., Chiu, Y. I., & You, W. (2007). The contribution of the Responsive Classroom approach on children's academic achievement: Results from a three year longitudinal study. *Journal of School Psychology, 45,* 401-421.

Rimm-Kaufman, S. E., Pianta, R. C., & Cox, M. J. (2000). Teachers' judgments of problems in the transition to kindergarten. *Early Childhood Research Quarterly, 15(2*), 147-166.

Rimm-Kaufman, S. E. & Sawyer, B. E. (2004). Primary-grade teachers' self efficacy beliefs, attitudes toward teaching, and discipline and teaching practice priorities in relation to the Responsive Classroom approach. *Elementary School Journal, 104*, 321 - 341.

Roberts, G., Vaughn, S., Fletcher, J., Stuebing, K., & Barth, A. (2013). Effects of a response-based, tiered framework for intervening with struggling readers in middle school. *Reading Research Quarterly, 48*, 237-254.

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the prospects study of elementary schools. *Teachers College Record, 104*, 1525-1567.

Sableski, M.-K. (2009). Scaffolding as an impetus for change when working with struggling readers. *Journal of Reading Education, 34(3), 30-37.*

Scarborough, H. S., & Parker, J. D. (2003). Matthew effects in children with learning disabilities: development of reading, IQ, and psychosocial problems from grade 2 to grade 8. *Annals of Dyslexia, 53,* 47-70.

Schafer, J. L. (1997) *Analysis of Incomplete Multivariate Data*. Chapman & Hall, London.

Schaefer, B. A., & McDermott, P. A. (1999). Learning behavior and intelligence as explanations for children's scholastic achievement. *Journal of School Psychology, 37,* 299-313.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147-177.

Schatschneider, C., Fletcher, J.M., Francis, D.J., Carlson, C.D., & Foorman, B.R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96,* 265-282. doi: 10.1037/0022-0663.96.2.265

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The high/scope Perry preschool study through age 40.* (Monographs of the High/Scope Educational Research Foundation, 14). Ypsilanti, MI: High/Scope Press.

Scott-Little, C., Kagan, S. L., & Frelow, V. S. (2006). Conceptualization of readiness and the content of early learning standards: The intersection of policy and research? *Early Childhood Research Quarterly, 21*, 153-173. doi: 10.1016/j.ecresq.2006.04.003

Singh, M. (2013). A longitudinal study of a state-wide reading assessment: The importance of early achievement and socio-demographic factors. *Educational Research and Evaluation: An International Journal on Theory and Practice, 19*(1), 4-18. doi: 10.1080/13803611.2012.737572

Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research, 75*(3), 417-453.

Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective Reading Programs for the Elementary Grades: A Best-Evidence Synthesis. *Review of Educational Research, 79*(4), 1391-1466.

Smokowski, P. R., Mann, E. A., Reynolds, A. J., & Fraser, M. W. (2004). Childhood risk and protective factors and late adolescent adjustment in inner city minority youth. *Children and Youth Services Review, 26*, 63-91. doi: 10.1016/j.childyouth.2003.11.003

Snow, Burns, & Griffin (Eds.). (1998) *Preventing reading difficulties in young children* [Electronic version]. Retrieved from: http://books.nap.edu/openbook.php?record_id=6023&page=1

Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first grade reading difficulties: The effects of oral language, literacy, and behavior skills. *Developmental Psychology, 41*, 225-234. doi: 10.1037/0012-1649.41.1.225

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.

StataCorp. (2009a). Stata 11 Multiple Imputation Data Reference Manual. College Station, TX: Stata Press.

StataCorp. (2009b). Stata 11 Survey Data Reference Manual. College Station, TX: Stata Press.

Staub, F. C., & Stern, E. (2002). The Nature of Teachers' Pedagocical Content Beliefs Matters for Students' Achievement Gains: Quasi-Experimental Evidence From Elementary Mathematics. *Journal of Educational Psychology, 94*(2), 344-355. doi: 10.3102/0162373712467080

Stipek, D. (2002). At what age should children enter kindergarten? A question for policy makers and parents. *Social Policy Report, 16,* 3 - 16.

Stipek, D. (2004). Teaching practices in kindergarten and first grade: Different strokes for different folks. *Early Childhood Research Quarterly, 19*, 548-568. doi: 10.1016/j.ecresq.2004.10.010

Stipek, D. & Byler, P. (2001). Academic achievement and social behaviors associated with age of entry into kindergarten. *Applied Developmental Psychology, 22*, 175-189.

Stipek, D., & Byler, P. (2004). The early childhood classroom observation measure. *Early Childhood Research Quarterly, 19*, 375-397.

Stipek, D. J., & Ryan, R. H. (1997). Economically disadvantaged preschoolers: Ready to learn but further to go. *Developmental Psychology, 33*, 711-723.

Stronge, J. H., Ward, T. J., Tucker, P. D., & Hindman, J.L. (2007). What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of Personnel Evaluation in Education, 20*, 165-184. doi: 10.1007/s11092-008-9053-z

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics, Fourth Edition*. Needham Heights, MA: Allyn & Bacon.

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. *The Elementary School Journal, 101*(2), 121-165.

Torgeson, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology, 40,* 7-26.

Torgeson, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., 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.

Tourangeau, K., Burke, J., Lê, T., Wan, S., Weant, M., Brown, E., … Walston, J. (2004). *Combined user's manual for the ECLS-K base-year public use data files and electronic codebooks* (NCES 2001-029r). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Nord, C., Lê, T., Pollack, J. M., & Atkins-Burnett, S. (2006). *Combined user's manual for the ECLS-K longitudinal kindergarten-fifth-grade public-use data files and electronic codebooks* (NCES 2006-032). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Tourangeau, K., Nord, C., Lê, T., Wan, S., Bose, & West, J. (2002). *Combined user's manual for the ECLS-K longitudinal kindergarten-first grade public-use data files and electronic codebooks* (NCES 2001-029r). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Vaughn, S., & Haager, D. (1994). The measurement and assessment of social skills. In G.R. Lyon, *Frames of reference for the assessment of learning disabilities* (pp. 555-570). Baltimore, MD: Paul Brookes.

Vernon-Feagans, L., Gallagher, K., Ginsberg, M.C., Amendum, S., Kainz, K., Rose, J., Burchinal, M. (2010). A diagnostic teaching intervention for classroom teachers: Helping struggling readers in early elementary school. *Learning Disabilities Research & Practice, 25*, 183-193.

Vitaro, F., Brendgen, M., Larose, S., & Tremblay, R. E. (2005). Kindergarten disruptive behaviors, protective factors, and educational achievement by early adulthood. *Journal of Educational Psychology, 97*, 617-629. doi: 10.1037/0022-0663.97.4.617

Vygotsky, L.S. (1978). *Mind in Society.* Cambridge, MA: Harvard University Press.

Walberg, H. J., & Tsai, S.-L. (1983). 'Matthew' effects in education. *American Educational Research Journal, 20*, 359-373.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*, 89-122.

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Educational Policy Analysis Archives, 10(12)*. Retrieved from http://epaa.asu.edu/epaa/v10n12/.

West, J., Denton, K., & Germino-Hausken, E. (2000). *America's kindergartners* (NCES Publication No. 2000-070R). Washington, DC: National Center for Education Statistics, U.S. Department of Education, Office of Educational Research and Improvement.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69,* 848-872.

Wiggan, G. (2007). Race, school achievement, and educational inequality: Toward a student-based inquiry perspective. *Review of Educational Research, 77*, 310-333. doi: 10.3102/003465430303947

Williford, A. P., Maier, M. F., Downer, J. T., Pianta, R. C., & Howes, C. (2013). Understanding how children's engagement and teachers' interactions combine to predict school readiness. *Journal of Applied Developmental Psychology, 34*, 299-309.

Wolter, K. M. (2007). Taylor Series Methods, pp. 226-271. Introduction to Variance Estimation. 2nd ed. New York: Springer.

Woodcock, R. W., McGrew, K. S., & Werder, J. K. (1994). Woodcock-McGrew-Werder mini-battery of achievement. Chicago: Riverside Publishing Company.

Wu, X., Anderson, R. C., Nguyen-Jahiel, K., & Miller, B. (2013). Enhancing Motivation and Engagement Through Collaborative Discussion. *Journal of Educational Psychology, 105*(3), 622-632.

Wu, J.-Y., Hughes, J. N., & Kwok, O.-M. (2010). Teacher-student relationship quality type in elementary grades: Effects on trajectories for achievement and engagement. *Journal of School Psychology, 48*, 357-387. doi: 10.1016/j.jsp.2010.06.004

Wu, W., West, S. G., & Hughes, J. N. (2008). Effect of retention in first grade on children's achievement trajectories over four years: A piecewise growth analysis using propensity score matching. *Journal of Educational Psychology, 100*, 727-740.

Xue, Y., & Meisels, S. J. (2004). Early literacy instruction and learning in kindergarten: Evidence from the Early Childhood Longitudinal Study – Kindergarten class of 1998-1999. *American Educational Research Journal, 41*, 191-229.

Yen, C.-J., Konold, T. R., & McDermott, P. A. (2004). Does learning behavior augment cognitive ability as an indicator of academic achievement? *Journal of School Psychology, 42,*157-169. doi: 10.1016.j.jsp.2003.12.001

Zabel, R. H., & Nigro, F. A. (2001). The influence of special education experience and gender of juvenile offenders on academic achievement scores in reading, language, and mathematics. *Behavioral Disorders, 26*(2), 164-172.

Zill, N., & West, J. (2001). *Entering kindergarten: A portrait of American children when they begin school: Findings from the Condition of Education 2000* (NCES Publication No. 2001-035). Washington, DC: National Center for Education Statistics, U.S. Department of Education, Office of Educational Research and Improvement.