5-2-2014

Examining the Relationship Amongst Fidelity of Implementation and Student Outcomes of a Tier I English Language Arts Curriculum for Adolescent Readers

Kelly J. McNamara
University of Connecticut, mcnamarak5@southernct.edu

Examining the Relationship Amongst Fidelity of Implementation and Student Outcomes of a Tier I English Language Arts Curriculum for Adolescent Readers

Kelly Jean McNamara

University of Connecticut

2014

As many scholars have noted, the educational system in the United States is entrenched in a literacy crisis (e.g., Boardman et al., 2008; Haynes, 2005; Moje et al., 2008; Torgesen et al., 2007). While this literacy crisis effects all ages and grade levels of students, a population of students that warrants particular concern is adolescent readers. Lee, Grigg, and Donahue (2007) found that approximately two-thirds of both eighth- and twelfth-graders read below proficiency and lack the reading skills needed to succeed in school. With the national movement towards a Response to Intervention (RTI) context of identifying and intervening with struggling learners, the Tier I general education classroom is the first line of defense in preventing and intervening with literacy difficulties. Proponents of RTI believe that comprehensive and coordinated instruction that is implemented with fidelity can improve outcomes for all students. However, fidelity has not been clearly defined in the literature, and has historically received less attention in the K-12 education literature than in other fields (such as the health field) (Ruiz-Primo, 2005; Summerfelt, 2003). The purpose of this study was to evaluate the correlation between teachers' fidelity to an English Language Arts curriculum and student outcomes on measures of reading achievement. An observational tool was constructed and used to record the fidelity of implementation middle school teachers of

English/Language Arts (ELA) classes employed over several observations of their teaching. Observational data was compared to student reading performance to examine relationships between fidelity and student reading performance. Results indicated relationships between dosage and reading performance in that students' reading scores increased as the time teachers spent delivering ELA instruction (dosage) increased. Additionally, a relationship between student engagement and independent practice indicated that the more time spent in independent practice in classrooms, the less engaged students were in instruction. Limitations to statistical power, the representativeness of the sample, and the observational tool created for the study are important to consider when interpreting the results of this study.

Examining the Relationship Amongst Fidelity of Implementation and Student Outcomes of a Tier I English Language Arts Curriculum for Adolescent Readers

Kelly Jean McNamara

B.A., Tufts University, 1997

M.A., Tufts University, 1999

C.A.G.S., Tufts University, 2001

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2014

APPROVAL PAGE

Doctor of Philosophy Dissertation

Examining the Relationship Amongst Fidelity of Implementation and Student Outcomes of a Tier I English Language Arts Curriculum for Adolescent Readers

Presented by

Kelly Jean McNamara, B.A., M.A., C.A.G.S.

Major Advisor _____
Melissa Bray

Associate Advisor _____
Michael Faggella-Luby

Associate Advisor _____
Lisa Sanetti

Associate Advisor _____
Thomas Kehle

Associate Advisor _____
Hariharan Swaminathan

University of Connecticut

2014

ACKNOWLEDGEMENTS

There are many people whom I would like to acknowledge for their professional and personal contributions in the development of this dissertation. First, I would like to thank my advisor Dr. Melissa Bray who has provided me with such wonderful support throughout this process. Her constant availability and sage guidance helped me to pass over innumerable hurdles. Furthermore, her kindness and compassion in both my personal and professional undertakings have been invaluable.

Dr. Michael Faggella-Luby also contributed significantly to my personal and professional development. There were many times along the path that I felt uncertain where to turn; in those times, Dr. Faggella-Luby was always there for me with thoughtful advice to help inform my decisions. When it came time to develop my dissertation proposal, our countless conversations about adolescent literacy were vital contributions to the development of concepts and methodology undertaken in this study.

Dr. Lisa Sanetti provided me with insightful commentary regarding the concept of fidelity of implementation, especially in regards to how these concepts were defined and measured in this study. Dr. Sanetti always challenged me to look more deeply and think more practically about how I defined concepts and ultimately interpreted my findings. This feedback helped shaped this study, and likely my future research undertakings.

Dr. Thomas Kehle has a broad and deep knowledge of the field of school psychology that has always inspired me to challenge myself to embrace life-long learning. Dr. Hariharan Swaminathan helped make the sometimes-daunting statistical analyses accessible; his calm demeanor and genuine interest in helping others grasp statistical concepts was immensely helpful throughout this process.

Dr. Kari Sassu was a source of constant support and assistance. From the beginning of my dissertation when I was challenged by subject recruitment, to the end when she offered insightful feedback regarding how to make points more clearly, I benefitted greatly from her kindness and guidance. Dr. Kenneth Pugh provided thoughtful feedback to the final product, and inspiring ideas for where to go next with this project.

Dr. Karen Rambo-Hernandez was instrumental in helping me approach the task of analyzing my data. Her astute insight into regression analysis and hierarchical linear modeling was a wonderful resource. Dr. Rambo-Hernandez patiently and kindly guided me through many of my analyses, and her careful feedback as I carried out the task of writing my results helped my data come to life in a meaningful and articulate fashion.

I also want to acknowledge my colleagues at Southern Connecticut State University and the cohorts of students that I have had the pleasure of serving as their instructor. They kept me motivated by continually reminding me why I had undertaken this project in the first place, and why I needed to finish.

Furthermore, I have been blessed to have many people in my life who have provided unconditional love and unflappable confidence in my pursuits and me. These include my entire family, but especially my parents and my brothers, and my dear friends Megan Wakely and Tammy Lerro. Finally, I owe a great deal of this accomplishment to my boyfriend, Chris Avery. He continually went out of his way to take things off my plate so that finishing this dissertation was a more attainable task, and perhaps more importantly, made sure that I had an endless supply of jellybeans to keep me going. He supported me unconditionally, and his encouragement was key to making it possible for

me to commit to this undertaking. To all of you, thank you for picking me up when I fell down, and believing in me when I forgot how to believe in myself. I am forever grateful that you are a part of my life.

This has been the most challenging undertaking of my life to date, but also the most rewarding. In the spirit of persevering over life's obstacles, I dedicate this to my brother Ryan, who continues to persevere in health and wellness. I also dedicate this dissertation to my nephews, who I hope are inspired to forever follow their passions, tirelessly, and with the spirit of a gracious and humble fighter.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction and Literature Review

As adolescents complete high school and ponder their future, there are two paths students predominantly consider: entering the work force immediately, or, pursuing higher education to gain specialized training before entering the work force. Research findings suggest, however, that today's adolescents are under-prepared for either of these paths, with almost 30% of high school graduates lacking the literacy skills necessary for higher education and approximately 40% lacking the literacy skills needed for the work force (ACT, 2005; Adelman, 2006; Hock & Deshler, 2003; Kamil, 2003). Astonishingly, regardless of the path chosen, one-quarter of our nation's adolescents cannot read material essential for daily living, such as road signs, newspapers, or bus schedules (National Center for Educational Statistics (NCES), 2005). These data paint a bleak landscape of adolescent literacy in America, suggesting that this country is facing something of a literacy crisis (e.g., Jacobs, 2008; Salinger, 2011). As the primary means through which adolescents receive instruction in literacy skills, schools (and teachers, specifically) are ideal agents for enacting change to address this literacy crisis.

The federal government, through the United States Department of Education (US DOE), closely monitors the educational status of the citizens it serves. When the US DOE was initially created in 1980 as a result of the passage of Public Law 96-88 in October, 1979 (U.S. Department of Education, Office of Communications and Outreach (DOE OCO), 2010), it was responsible for four major activities, including the collection of data and research on America's schools. In researching American schools, the US DOE collects information regarding the educational attainment of our nation's students in

many important areas, including literacy. Since its inception, data have indicated increasing alarm over the literacy status of American adolescents; in fact, many scholars have proclaimed that we are facing an adolescent literacy "crisis" (e.g., Jacobs, 2008; Salinger, 2011). It is startling to realize that in over twenty years (comparison of National Assessment of Educational Progress (NAEP) data from 2007 and 1985), adolescent literacy rates have remained essentially unchanged (Jacobs, 2008). When one considers that literacy gains for adolescents were unchanged or increased in an insignificant manner from 1971 to the 1985 report (Jacobs, 2008), it is clear to see that literacy gains among adolescent readers have been essentially non-existent in the last 40 years; certainly, adolescent literacy is in crisis.

The US DOE has responded to the adolescent literacy crisis in many ways over the years, typically by legislating mandates intended to improve literacy skills among adolescent readers. Two of the most notable and impactful pieces of federal legislation to address the literacy issues facing our nation's adolescents include the No Child left Behind (NCLB) Act of 2001 (the reauthorization of the Elementary and Secondary Education Act) and the Individuals with Disabilities Education Improvement Act (IDEIA) of 2004 (reauthorization of the Individuals with Disabilities Education Act of 1990 (originally the Education for All Handicapped Children Act of 1975- Public Law 94-142)). While NCLB had four main areas of focus, two areas paid direct attention to literacy: increased accountability for states, school districts, and schools; and, a stronger emphasis on reading (NCLB Executive Summary, 2002). The latter stronger emphasis on reading is admittedly predominantly focused upon the improvement of early reading skills among early elementary grades; however, this emphasis stems from the knowledge

that students are graduating our public schools with poor literacy skills (e.g., ACT, 2005; Adelman, 2006; Hock & Deshler, 2003; Kamil, 2003; Salinger, 2011), and an effective and humane way to address this deficit in adolescence is to prevent it from ever occurring while students are young and still developing their literacy skills (Albee & Gullotta, 1997; Burns, Griffin, & Snow, 1999). The former focus on increased accountability forces states, districts, and schools to collect and examine data on student academic progress, and then analyze those data and use it to make better decisions about allocating resources to address learning needs. Thus, both of these elements of NCLB lay a foundation for addressing the literacy crisis currently experienced by adolescents in our nation.

The second piece of legislation to have a direct impact on the adolescent literacy crisis is IDEIA (IDEA 2004). While this piece of legislation is primarily intended to help identify and safeguard students with disabilities within our nation's public schools, the methods advocated in this legislation for identifying students with disabilities has created an opportunity for ***all*** struggling learners to access better intervening services (see Connecticut State Department of Education, Bureau of School and District Improvement (CSDE), 2008). Citing years of research, the IDEIA legislation promotes the use of a Response to Intervention (RTI) model to identify and provide intervention to students at-risk for not meeting educational performance standards. While a single authoritative definition of RTI does not exist (NICHCY, n.d.), the general RTI framework includes: evidence-based instruction, tiered instruction of increasing intensity, screening children within the general curriculum, close monitoring of student progress, and informed

decision making regarding next steps for individual students (http://nichcy.org/schools-administrators/rti#what).

All five of the above elements of RTI contribute in a meaningful manner in responding to our nation's adolescent literacy crisis. First, evidence-based instruction ensures that all teachers deliver instructional practices that have an evidence-base from which their use can be supported. The research regarding quality instructional practices is relatively well-defined (Colvin, Flannery, Sugai, & Monegan, 2009), and general consensus exists as to what quality instruction should entail in adolescent ELA classrooms. Using policy documents such as *Reading Next* (Biancarosa & Snow, 2006), and reviews of adolescent literacy research (e.g., Faggella-Luby, Ware, & Capozzoli, 2009), we understand the essential components and practices that need to be included in all classrooms where quality adolescent literacy instruction is enacted. Ensuring this access to high-quality instruction helps ensure that students who come to be identified as struggling to meet the demands of the literacy curriculum are not falling short due to instruction grounded in poor or unfounded instructional practices.

Second, tiered instruction of increasing intensity provides a systematic method of ensuring that students' instructional and learning needs are met. While there are variations, the tiered system is generally conceptualized as a three-tiered format wherein ***all*** students are included in Tier I instruction, and some students ***also*** receive additional, more intensive instructional support and intervention at Tier II and/or Tier III, based upon information gained from universal screenings and progress-monitoring of students within the tiered interventions (both universal screening and progress monitoring are described below). In Tier I, all students are provided with a coordinated, research-based

instructional core that has been designed to ensure that students successfully master the literacy curriculum. In fact, literacy curriculums delivered at Tier I are generally considered effective when approximately 80% of students receiving the curriculum meet benchmark literacy goals in universal screenings (CSDE, 2008, p. 30). Thus, with a comprehensive and coordinated literacy curriculum that is delivered with quality, evidence-based practices at Tier I, most students should meet literacy expectations. The 20% or so who do not meet these benchmarks, are identified as at-risk for failing to meet such benchmarks and provided instruction of increased intensity at Tier II. These interventions can vary in duration, dosage, skill-level of intervention implementer, etc., but what distinguishes them from Tier I instruction is that the instruction is more intensive (perhaps for a longer period of time, perhaps delivered by a more highly-trained individual, etc.) and *in addition to* the quality instruction delivered at Tier I. If data collected to monitor progress of the Tier II intervention demonstrate that the at-risk student is not responding to the Tier II intervention, the intervention can be intensified again at Tier III. By having these progressively more intense levels of quality instruction available to ***all students*** who have been identified as at-risk for not meeting benchmark goals, schools can more quickly identify and intervene with students who struggle. In the past, schools largely used a "wait to fail" instructional model, only providing intervention to students when they demonstrated a significant discrepancy between their cognitive aptitude and their academic achievement (Fuchs, Mock, Morgan, & Young, 2003). Thus, students often struggled for years before their skills were "low enough" to warrant intervention. With the RTI model, a student simply has to fail to meet a benchmark standard at a regularly scheduled screening to trigger more intensive intervention. By

intervening more quickly, we are now able to intervene at the first sign of struggle, rather than waiting for clear signs of failure. This more immediate approach to intervening with adolescent literacy struggles promotes a proactive response to improving literacy rates among our nation's adolescent readers.

Third, screening of students within the general curriculum ensures that ***all*** students are assessed to determine if they are at-risk for not reaching established performance benchmarks (typically set by individual school districts) in a variety of areas, including literacy. These screenings are considered ***universal*** because all students are screened, thereby reducing the likelihood that a student will fall through the cracks when identifying learning struggles. Typically schools administer these screenings at pre-established intervals several times throughout the school year, thus providing several opportunities to identify students who are not at the established performance benchmark, and thus at-risk for educational difficulties. These screenings typically begin in kindergarten and are intended to extend into the high school years; as such, students are continually monitored throughout their education to determine their level of risk.

Fourth, close monitoring of student progress ensures that the delivered intervention, whether at Tier II or Tier III, is having the desired and appropriate effect upon the student to whom it is delivered. By taking frequent (often weekly) data, the interventionist can determine if the selected intervention, and schedule with which it is delivered (including dosage and other relevant components of the intervention), is effective in improving the student's literacy skills and reducing their risk for academic failure. When these data indicate that the selected intervention is not having the desired

effect upon the student, school staff can make informed decisions for altering the intervention so that it is more effective.

This brings us to the final fifth element of RTI, informed decision-making regarding next steps for individual students. Using data from universal screening and timely progress monitoring assessments, school staff can make *informed* decisions about how to intensify interventions or remove scaffolds on interventions. As such, students can then be provided with the appropriate instruction at the appropriate Tier to increase literacy for all students.

The Importance of powerful curriculums at Tier I in the context of RTI. Given that Tier I instruction is a "first-line" defense against literacy struggles, and that we generally expect that 80% of students receiving the Tier I curriculum will respond and reach expected literacy benchmarks through Tier I instruction alone, it is imperative these curriculums be comprehensive and coordinated (CSDE, 2008). The curriculum must be comprehensive, addressing the range of academic literacy competencies at each grade level. To best leverage instructional time, these curriculums also need to be coordinated, ensuring the logical, developmental progression of skills across grades. But even with a comprehensive, coordinated, research-based curriculum, if it is not implemented with quality and fidelity, it will likely fail to produce the desired outcome of ensuring that the majority of students who receive the curriculum (approximately 80% of students) meet established benchmark literacy goals. As such, ensuring the fidelity of the delivered curriculum is a critical element to addressing our nation's literacy crisis.

What is fidelity of implementation? Given the importance of this concept in ensuring effective literacy instructional practices within our schools, it is essential that a common

understanding of the term be established. When examining the research related to fidelity of implementation, one quickly finds that the literature is still emerging, and historically has not been well developed. The emergence of this concept in the literature struggled to take off, in part, due to the fact that this concept is known by many terms (e.g., treatment integrity, procedural reliability, etc.) and is not consistently reported when discussing treatment or research outcomes. In their introduction to the special series "Toward Developing a Science of Treatment Integrity" in *School Psychology Review*, Sanetti & Kratochwill (2009) proposed the following broad working definition of implementation fidelity, which they refer to as treatment integrity: treatment integrity is "the extent to which essential intervention components are delivered in a comprehensive and consistent manner by an interventionist trained to deliver the intervention" (p. 448).

Building off the above conceptual definition of implementation fidelity, one can start to elaborate upon what essential elements need to be considered when trying to create an operational definition of this construct. Examination of the many common definitions of implementation fidelity (e.g., Dane & Schneider, 1998; Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005; Jones, Clarke, & Power, 2008; Noell, 2008; Power et al., 2005; Waltz, Addis, Koerner, & Jacobson, 1993) reveals when attempting to define the term, four components consistently emerge as foundational: the content of the intervention, the process of the intervention, the quantity of the intervention, and the quality of the intervention (Sanetti & Kratochwill, 2009). The content of the intervention is generally *what* was delivered; that is, what steps were followed to deliver the intervention to the student(s). Once these steps are defined, implementation fidelity is typically assessed, in part, by assessing how closely the defined steps were followed, or

how closely the interventionist *adhered* to the steps of the intervention (Waltz et al., 1993).

A second foundational component of fidelity of implementation, the process of the intervention, generally refers to *how* the intervention was delivered (Sanetti & Kratochwill, 2009). This process typically includes things such as who delivered the intervention, training given to the interventionist, and other important aspects of the process of delivering an intervention.

The quantity of the intervention is generally *how much* of the intervention was delivered by the interventionist (Sanetti & Kratochwill, 2009). Not to be confused with adherence, quantity is not necessarily concerned with how many steps were delivered, but rather, how much time was spent delivering the intervention (Dane & Schneider, 1998; Waltz et al., 1993). As such, quantity can be considered the *dosage* of the intervention delivered (time of individual sessions, as well as total time spent in the delivery of the intervention).

Finally, quality is generally conceptualized as *how well* an intervention was delivered. As such, when considering "how well" an intervention is delivered, one must consider the basic instructional principles and literature regarding quality instructional practices that are specific to the discipline; in this case, specific to literacy. Given these definitions, one can see that the elements of fidelity of implementation that pertain to adherence and dosage are easier to define and observe; as such, these aspects of fidelity of implementation are more commonly reported, whereas the equally important elements of process and quality are sometimes overlooked (e.g., Noell et al., 2005). Thus, we need to take care to fully define ***all*** elements of fidelity of implementation in all of our

scholarly and professional undertakings to ensure that the literature continues to grow and advance our understanding of this critically valuable element of core curriculums and the delivery of high-quality instruction in adolescent literacy classrooms.

Lack of attention to fidelity of implementation in K-12 literature. In addition to these problems with defining implementation fidelity (or perhaps, even, because of these problems), the concept of implementation fidelity has received comparatively less attention in the K-12 literature than in other fields, or than might be expected given the importance of this concept in determining research and intervention outcomes (Dobson & Cook, 1980; NRC, 2004; U. S. Department of Education, 2006). When one does not attend to, or report upon, the fidelity of implementation that was achieved with an intervention, results of the intervention must be interpreted with caution, as one cannot conclude whether the intervention truly yielded the observed results, or whether some other factor contributed to the intervention's success or failure. That is, without an understanding of how closely the delivery of the intervention followed its prescribed course (the fidelity with which it was implemented), it is possible that some other factor, such as an overlooked step in the treatment, a less highly trained interventionist that may have substituted for the interventionist on a given day, a change in schedule, or a variety of other factors, truly contributed to the results derived from the intervention. Thus, attending to the fidelity of implementation of all interventions is paramount to an intervention, as it provides validation that the intervention itself truly produced the observed results. Furthermore, research indicates that high levels of fidelity of implementation are associated with better outcomes (e.g., Durlak & DuPre, 2008; O'Donnell, 2008); however, the converse has not been indicated (e.g., Noell, 2008;

Sterling-Turner, Watson, & Moore., 2002). When fidelity of implementation is low, results observed cannot be attributed to the intervention, and data suggesting that it succeeded or failed to have an effect become insignificant at a practical level (Yeaton & Sechrest, 1981). As such, attending to the fidelity with which we implement interventions will likely enable us to achieve better outcomes with the evidence-based interventions we undertake to solve a problem. In education, the problems we attempt to solve are largely related to building student capacity to learn and become better citizens, so attending to fidelity of implementation to improve the outcomes of these problems is imperative.

Fidelity of Implementation in School Settings. Given the above-noted lack of attention to fidelity of implementation in the K-12 literature, it is not entirely surprising that this concept has historically rarely been reported in large-scale education studies that examine the effectiveness of K-12 core curriculum interventions (Dobson & Cook, 1980; NRC, 2004; U. S. Department of Education, 2006). Given the importance of the core literacy curriculum in addressing the current literacy crisis among adolescents in our nation, coupled with what we know about how fidelity of implementation can increase or limit the effects of intervention on outcomes, devoting increased attention to the fidelity with which core literacy curriculums are implemented is no longer something that can be ignored. Unfortunately, due to this lack of attention in the literature, there are currently not many studies to guide researchers in understanding how fidelity of implementation of core curriculum can be measured (O'Donnell, 2008). Furthermore, we also lack a clear understanding of how fidelity of implementation to core curriculums are related to outcomes (O'Donnell, 2008). If a strong, comprehensive, coordinated core literacy

curriculum is to be a front-line defense against poor literacy, we must learn effective ways to measure and report outcomes of core curriculum with specific attention given to the fidelity with which such curricula are implemented.

One of the difficulties in studying fidelity of implementation to core curricula is that there has been minimal development of viable and efficient measures of fidelity of implementation that have adequate psychometric properties (Sanetti & Kratochwill, 2009). Measures that do exist tend to be specific to an intervention, and as such, have limited ability to be applied to the variety of interventions that are implemented in schools, including core curricula. The few studies in the literature that have offered some insight into measuring and accounting for fidelity of implementation often report beginning the process of developing criteria for determining fidelity of implementation with a curriculum profile or analysis (O'Donnell, 2008). This analysis would define the critical components of the intervention/curriculum with the researcher or developer of the curriculum then outlining ranges of acceptable variation within these parameters of implementation (Songer & Gotwals, 2005). This process typically yields a component checklist, which can then be used to record fidelity of implementation to the curriculum. When developing such checklists, one must be careful to give equal attention to all of the components of fidelity of implementation, including content, process, quantity, and quality. Mowbray, Holter, Teague, and Bybee (2003) stated that process criteria may be more difficult to measure reliably, but they may also be more significant in terms of program effects (O'Donnell, 2008).

Assessing fidelity of implementation to core curricula. To effectively measure fidelity of implementation to core curricula, one must begin with a clear definition of the

curriculum one intends to measure. Thus, it is important to understand how curriculum is defined. Remillard (2005) has provided a definition that is helpful when considering core curricula; he defines it as the process by which "individual teachers interact with, draw on, refer to, and are influenced by the material resources designed to guide instruction" (p. 212). Curriculum-in-use, however, appears to be a separate entity, one that is viewed as that which is implemented by the teachers and not necessarily identical to the written curriculum (Munby & Russell, 1990; Shkedi, 1998). Given these variations in how one defines curriculum, the challenge in measuring fidelity of implementation to core curricula begins with determining how to distinguish between the effects caused by the materials and the effects caused by the teachers' interactions with the materials (O'Donnell, 2008). In this vein, it is clear that measuring the content (adherence) and quantity (dosage) of the curriculum are important to measuring fidelity of implementation of the curriculum, as that will provide the information related to the effects caused by the curriculum materials themselves. Equally important, but arguably more difficult to assess, is the quality of the implementation and the degree of program differentiation, as this will provide insight into the effect of the teachers' interactions with the curriculum materials. Participant responsiveness may provide a unique window into the interaction of these equally important elements of fidelity of implementation, as it will likely be the product of both the curricular materials as well as the teachers' interactions with these curricular materials. Thus, researchers must assess adherence to the curriculum, dosage of the curriculum, quality of implementation of the curriculum, participant responsiveness to the curriculum, and degree of program differentiation, where applicable.

Adherence. As noted earlier, adherence refers to how closely the steps of an intervention are followed. When relating the concept of adherence to a school setting in which one is interested in observing fidelity of implementation to a core curriculum, the curriculum itself would be considered the "intervention." Thus, in this context, researchers need to determine how closely the steps of the curriculum are followed, and to do so, one must identify the steps of the curriculum. Since school districts generally establish their own curricula (Hattie, 2005), there is tremendous variance in not only what the curricula entails, but also in how well the curricula is defined (Hattie, 2005). Thus, in order to reliably measure a teacher's adherence to a core curriculum, it is essential that the school district have a written description of the curriculum that details the essential elements for implementation. Components such as a scope and sequence and pacing guide are helpful features that can elucidate *what* the essential elements, or steps, of the curriculum are, as well as *when* the elements should be taught during the school year. With this information, one could consider a teacher to be in adherence with the school district's written curriculum if they teach *what* the curriculum describes *at the designated time* in the school year.

Dosage. Dosage refers to how much of the intervention was delivered, or the quantity of the intervention delivered (Sanetti & Kratochwill, 2009). As with adherence, when relating the concept of dosage to a school setting, the curriculum being observed is the intervention. Thus, dosage, in this context, refers to how much of the written curriculum was delivered by the teacher. This can be thought of over the course of a school year (did the teacher implement the entire scope of the written curriculum?) or over the course of a single day (did the teacher implement the entire lesson?). For studies

concerned with teacher fidelity of implementation over the course of an entire school year, it would be essential to document *what* was covered on a regular basis to ensure that all of the essential elements were delivered by the end of the school year. For studies concerned with day-to-day implementation of curriculum over shorter spans of time, it is essential to document *how much* of the allocated instructional time was spent delivering the curriculum. For example, if a school schedules class periods for 60 minutes daily, then an English/Language Arts (ELA) class would have 60 minutes available daily for ELA instruction. To assess the dosage of a curriculum on a daily basis, a researcher would need to determine how much of the allocated 60 minutes was spent in instruction (as opposed to time spent in transition, time spent unengaged in instruction, etc.).

Participant responsiveness. The element of fidelity of implementation considered participant responsiveness is often defined as how participants respond to, or are engaged by, an intervention (Carroll et al., 2007). Participant responsiveness is an important element of fidelity of implementation to consider. Given that even a high quality curriculum can fail to produce positive results if the students do not interact with the curriculum, it is important that students engage in learning behaviors that will allow them to benefit from the high quality curriculum being offered by the school district (Carroll et al., 2007). When considering participant responsiveness in the context of a core curriculum implemented by a teacher in a school setting, the participants are the students receiving the teacher's instruction. To measure the student's response to the curriculum, one can measure their achievement in the curricular area (response to the curriculum) as well as their engagement in the curriculum. In keeping with the example of a school ELA curriculum, response to the curriculum could be assessed through gains

in achievement on measures of literacy, as well as through observation of the students' engagement in lessons delivered by teachers. To measure the former component of responsiveness, one could administer tests, quizzes, and other assessments to determine the learning, or response, to the curriculum. To measure this latter engagement component of responsiveness, one would need to observe the actions, behaviors, and interactions students have with the teacher, lesson content, and each other over the course of lessons implemented from the written curriculum. These types of observations could then be used to determine the amount of instructional time in which students are engaged in these learning actions, behaviors, and interactions.

Program differentiation. Program differentiation refers to determining the unique components of a program, or those elements that are essential for the success of any program or intervention (Carroll et al., 2007). This aspect of fidelity of implementation is important to measure in instances where a curriculum is adopted that was developed for a different population of students, different grade level, different amount of time allocated to the delivery of instruction, different ELA content, etc. than that which is represented by the school in which the curriculum will be implemented. When a written curriculum needs to be adapted to meet the needs of a specific school or school population, it is crucial to ensure that all of these essential elements are present in the adopted version of the curriculum, indicate all adaptations made, and assess how these adaptations impact the implementation of the curriculum. For example, if an essential component is difficult to implement in a new setting, the absence of this essential element may explain any lack of success of the curriculum rather than a simple failure of the program (in its original form with all essential elements present).

Quality of implementation. This fifth element of fidelity of implementation generally refers to how well or the manner in which a program or intervention is implemented (Carroll et al., 2007; Sanetti & Kratochwill, 2009). In the context of a school ELA classroom, quality of implementation would be considered the quality of the instruction a teacher uses to deliver the curriculum. As noted earlier, the literature regarding quality instructional practices is relatively well-defined (Colvin, Flannery, Sugai, & Monegan, 2009), and general consensus exists as to what quality instruction should entail in high school ELA classrooms. To demonstrate quality instruction, teachers in these classrooms must be *engaged* in *teaching functions* that consistently include *indicators of quality teaching* that provide students with appropriate *feedback* while ensuring an effective learning arrangement that keeps *transition time* to a minimum and *students engaged*.

Transition Time. In the context of a high school LEA classroom, transition time refers to time spent moving from one place, activity, lesson, etc. to another. Transition time is class time that is not devoted to instruction, often because instruction has not yet begun (students have not yet settled into class) or because the teacher is moving the students from one activity to another. Transition time is not necessarily *unengaged* time, as the teacher may still be engaged, in that they are directing students in their transition. Since transition time takes away from time that can be devoted to instruction within classroom settings, it is important that transition time be minimized. In fact, research demonstrates that time spent actively engaged in instruction contributes to student achievement (Walberg, 1986). Research in elementary classrooms has indicated that transition time accounts for almost 15% of classroom instructional time (Gump, 1967).

High school classrooms differ from elementary classrooms, however, in that students are typically in classrooms for shorter periods of time (45-90 minutes in content-specific high school classroom as opposed to about 6 hours of self-contained instruction across all content areas in elementary classrooms). Given this difference, many have speculated that transitions in high school classrooms should occur less frequently and take less time than those that occur in elementary classrooms (Doyle, 1986) due to the reduced amount of available instructional time and content-specific focus of these classrooms.

Engagement. Engagement refers to the actions, behaviors, and interactions that the teacher/student(s) partake in to demonstrate their involvement with instruction. As noted above, it is crucial that teachers maximize the amount of allocated class time devoted to instruction to provide students with the best opportunity to benefit from instruction and attain high levels of achievement. Thus, teachers need to decrease behaviors that detract from their ability to remain engaged in instruction, such as checking emails, answering phone calls, talking with another adult, etc. Time in which teachers are not engaged in instruction is instructional time lost.

In addition to maximizing teacher engagement in instruction, it is also necessary to maximize students' engagement in the instruction provided by the teacher. Even in classrooms where teachers have maximized instructional time and consistently provide high-quality instruction, if students are not engaged in that instruction, they are not in a position to benefit from such instruction, often resulting in limited achievement.

While there is a clear case for documenting engagement in instruction (both on the part of the teacher and the students), engagement is often a difficult construct to measure through observation (Fredricks, Blumenfeld, & Paris, 2004) due to the

differences in defining engagement as well as the limitations involved with recording a construct that can be difficult to observe. Definitions of engagement are often intertwined with the concept of motivation (Fredricks, Blumenfeld, & Paris, 2004), making it difficult to disentangle the concepts and define and measure them separately. Furthermore, some have defined engagement as a construct with many layers, including behavioral, cognitive, and emotional components (e.g., Fredricks, Blumenfeld, & Paris, 2004). Behavioral engagement is more readily measured through observation, as this element of engagement refers to the participation in an activity (the work one does and the rules one follows to complete an activity). Cognitive and emotional engagement, however, are less readily measured through observation as they refer to investment one puts into an activity and the emotional connection one has with an activity, respectively, that cultivate a desire to perform the activity. As such, behavioral engagement can be tied to one's on-task behavior, which is frequently measured through observation. While observational methods of behavioral engagement are not without controversy (in that individuals are often judged to be engaged when they are in fact disengaged (daydreaming despite looking in the direction of instruction) and conversely judged to be disengaged when they are in fact engaged (doodling while relating new information to what they already know), such methods are commonly used to document this crucial element of fidelity of instruction.

Teaching Function. Teaching function refers to behaviors and procedures teachers engage in to deliver instruction to and assess learning of students. Rosenshine and Stevens' seminal 1986 synthesis of research regarding elements of the practice of

effective teaching resulted in six teaching functions that describe a best practice approach to instruction. These six teaching functions are:

1. Review
2. Presentation
3. Guided Practice
4. Corrections and feedback
5. Independent practice
6. Weekly and Monthly reviews

Five of these six functions originally described by Rosenshine and Stevens (1986) have been adapted into the six discrete categories of teaching function defined below for this study. The sixth teaching function described by Rosenshine and Stevens (1986), weekly and monthly reviews, has not been included as it is beyond the scope of the current study.

Review. When teachers are engaged in review, they are reviewing previous learning. This behavior of reviewing and summarizing the important elements of a lesson has been demonstrated to be associated with improved student achievement (Armento, 1976; Wright & Nuthall, 1970; as cited in Cornett, 2010). Associated with review is the concept of preview, or describing what students are about to learn. This is often referred to as a advance organizer, a vehicle by which educators can overview the learning that will occur in the lesson. Activities such as reviewing the homework or assignments assigned the previous day, reviewing a previously taught strategy or concept, or reviewing previously read or discussed text, are all examples of teaching functions that are considered review behaviors. Activities such as describing a text students will read,

describing a strategy or concept students will learn, describing activities students will engage in, stating learning goals of the lesson, etc. are all examples of advanced organizer behaviors.

Presentation. This teaching function involves the teacher engaging in delivery of lesson content for the intended lesson. This instruction should be devoted to the delivery of new learning, whether the entire presentation is new, or the teacher is presenting a new aspect of something previously learned. The teacher may communicate this new material to the students through demonstration, lecture, modeling, providing directions, providing examples and non-examples, asking students frequent questions to check their understanding, etc. When presenting new information, teachers should strive to present all skills, strategies, concepts, tasks, etc. in small steps to help ensure students will not be overwhelmed by the new material.

Guided Practice. Guided practice involves the teacher explicitly guiding the students through learning, typically by actively modeling and assisting students in practicing the new learning. Teachers may demonstrate engagement in this teaching function by controlling the complexity of tasks assigned to students (such as presenting multiple versions of assignments to students with each version targeted at the student's individualized level of instruction), assisting students in organizing knowledge, providing prompts to assist students in completing work correctly, monitoring student progress by asking students questions, circulating among students, checking work in progress, offering feedback, etc.

Independent Practice. This teaching function is defined by the teacher directing the students to work independently. Independent practice helps students to accumulate,

develop, and master skills, strategies, concepts, etc. presented by the teacher during presentation and guided practice. Activities teachers may ask students to complete independently include reading an assigned passage from the current reading selection, reading for pleasure, reading to gain information, working independently on a worksheet or other quick, in-class assignment, working independently on an extended writing assignment, beginning homework, etc. (all without assistance from any instructor or peer).

A subset of independent practice involves students working collaboratively in small peer groups (of at least 2 students), as directed by the teacher. Peer pair groups at times may be larger than two students (e.g., if there is an odd number of students and one or more groups is a group of three students working collaboratively). Research indicates that instructional arrangements where peers learn with and from each other can increase academic performance (Cushing & Kennedy, 1997; Harris, Marchand-Martella, & Martella, 2000; Maheady, Sacca, & Harper, 1987; Maheady, Sacca, & Harper, 1988). As such, *peer practice* is a separate yet important element of independent practice to consider within this teaching function.

Corrections and Feedback. When engaged in corrective feedback, the teacher is providing feedback to student responses. Whenever students respond to instruction, whether in oral, written, visual, or some other form, students should be given feedback as to how well their response addresses the question or assignment that triggered the student response. In reviewing 10+ years of research into the effects of schooling on student achievement, Hattie (1999) found that "the most powerful single moderator that enhances achievement is feedback" (p. 9). When students respond incorrectly, teachers should

immediately correct errors by identifying the error and providing the correct response (e.g., Archer & Hughes, 2011). Error correction should always be carried out in a positive manner that builds students' self-efficacy rather than creating a sense of failure (Guthrie, Anderson, Alao, & Rinehart, 1999). Positive reinforcement should be provided when students offer correct responses; that is, it is important to provide feedback to both correct and incorrect student responses.

Formative progress monitoring is a type of feedback the teacher engages in to help drive instruction; the purpose is to inform the teacher's instruction. Formative assessments, like quizzes, are often short, and frequently consist of curriculum-based assessments and measures. Circulating around the room or meeting individually with students or with small groups of students to check work in-progress is another method of engaging in feedback that can help drive instruction.

Formal Assessment. Tests and quizzes are types of formal assessments that measure student knowledge and assist teachers in assigning grades to students for their performance. Formal assessments provide teachers with a sense of what students have mastered in terms of the instruction they were provided. Formal assessments can be short (quiz) or more lengthy (test) in terms of the amount of instructional time they require to administer them to the students.

Quality Teaching Indicators.

Content Knowledge Instruction. Content knowledge instruction refers to instruction devoted to increasing students' knowledge of the content area. Teachers engage in content knowledge instruction by building upon and clarifying what students already know, and introducing new declarative, procedural, and conditional knowledge.

Teaching behaviors often encompassed in descriptions of content knowledge instruction include: building/clarifying background knowledge, declarative knowledge, procedural knowledge, and conditional knowledge (e.g., Archer & Hughes, 2011; Coyne et al., 2009; Ehren, 2005; Mastropieri, Scruggs, & Graetz, 2003; Rosenshine, 1995). Background knowledge is the information students already have about a particular topic, while declarative knowledge is factual or conceptual information, procedural knowledge is information about how something is done, and conceptual knowledge is information about when to use skills, strategies, or other information. Ehren (2005) noted that the importance of background knowledge is acknowledged and is especially relevant to the instruction and intervention practices with students who lack either experiential knowledge or knowledge gained from reading, or who need explicit instruction in how to use their past experiences to comprehend text (Ehren, 2005). Research tells us that good readers draw from, compare, and integrate their prior knowledge with material in the text (Duke & Pearson, 2002). This practice of engaging prior knowledge and experience prior to reading has positive effects on story understanding (Duke & Pearson) in that reading comprehension results when readers match what they already know (their schema) with new information and ideas in a text (Buehl, 2009). Proficient readers activate prior knowledge before, during, and after reading and they constantly evaluate how a text enhances or alters their previous understandings (Buehl, 2009). Building upon this background knowledge, students are positioned to more easily acquire new declarative, procedural, and conditional knowledge that will facilitate new learning in the content area.

Word Reading Instruction. Word reading instruction refers to teaching in which teachers provide students with instruction regarding how to read, understand, and use unknown words. Teaching behaviors often encompassed in descriptions of word reading instruction include: building independence in understanding unknown words through instruction in multi-syllabic word reading strategies and use of available reference materials, providing direct instruction in domain-specific and all-purpose academic vocabulary, and providing students repeated exposure to new words in multiple contexts (e.g., Baumann et al., 2003; Baumann et al., 2002; Bear, Invernizzi, Templeton, & Johnston, 2007; Curtis, 2004; Ehren, 2005; Henry, 2003; Jenkins, Matlock, & Slocum, 1989; Moats, 2001; Snow, Porche, Tabors, & Harris, 2007; Stahl & Fairbanks, 1986; Templeton, 2004; Wexler, Edmonds, & Vaughn, 2008). Since there are simply too many words in the English language to learn through direct instruction (Graves, Juel, & Graves, 2004), it is essential that student receive instruction in word reading/vocabulary strategies that provide them skills to acquire new word/vocabulary understandings independently at all grade levels. By the time students reach high school, the texts to which they are exposed are increasingly content-focused, containing many domain-specific words that often are not part of the student's oral vocabulary (Kamil et al., 2008). When students cannot match a new word to a known word within their accumulated oral vocabulary, students need tools to access the meaning in these unknown words. Research has demonstrated that such instruction improves students' ability to acquire domain-specific vocabulary commonly encountered in their content-area textbooks (Baumann, Edwards, Boland, Olejnik, & Kame-enui, 2003). Direct instruction in the spelling and meaning of unknown words, along with instruction in syllabication, morphology, strategies for

decoding multisyllabic words, and the use of reference materials (such as dictionaries, glossaries, thesaurus, etc.) all have been shown to improve students' ability to independently acquire vocabulary knowledge (Bear, Invernizzi, Templeton, & Johnston, 2007; NICHD, 2000).

Comprehension Strategy Instruction. Comprehension strategy instruction refers to teaching in which teachers provide students instruction regarding how to use specific reading strategies to improve comprehension of text. Teaching behaviors often encompassed in descriptions of comprehension strategy instruction include: teaching goal-specific strategies, such as comprehension monitoring, use of graphic organizers, summarizing/paraphrasing, question asking/generating, knowledge of text structures and features, inferencing, and finding main ideas; teaching packaged strategies; and, providing procedural prompts to use strategies (e.g., Biancarosa & Snow, 2006; Coyne et al., 2009; Ehren, 2005; Malone & Mastropieri, 1992; Mastropieri, Scruggs, & Graetz, 2003; National Assessment Governing Board, 2007; National Institute of Child Health and Human Development [NICHD], 2000; Rosenshine, 1995; Snowling & Hulme, 2005; Swanson, 1999; Swanson & Deshler, 2003). Comprehension monitoring, the use of graphic organizers, and instruction in the skills of summarizing/paraphrasing were all supported as evidence-based practices to improve students' reading comprehension by the National Reading Panel's 2000 publication *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction*, which identified practices that had an evidence-base to support their use in literacy instruction. Comprehension monitoring is a strategy that teaches students to understand and be aware of their own comprehension of text as they read; that

is, it strives to help students gain insight into that which they understand and that which is unclear when they read. Graphic organizers are instructional tools that transform words and ideas in text into graphic representations, providing a visual avenue for students to help construct meaning from text they read. Summarizing is the act of taking a large amount of text and reducing it to main/key points, while paraphrasing is the act of rewriting words or ideas in one's own words without altering the meaning of the original passage. Both are useful strategies for helping students increase their comprehension of text in that they assist students in focusing on what is important in a text, and distinguishing important information from less important information.

Mastropieri, Scruggs, and Graetz's 2003 review of comprehension strategies found that simple comprehension strategies that require students to ask/generate and answer questions about text while they are reading had a strong effect on students who previously did not appear to use comprehension strategies, and that self-questioning demonstrated the largest effect size of all comprehension strategies reviewed. Thus question asking/generating, particularly self-questioning, has a demonstrated effect in assisting students in building their comprehension of text. Knowledge of text structures and features helps students make connections between the ideas they read in text, in addition to assisting students in distinguishing important from less important content (Saenz & Fuchs, 2002). Research supports the use of instruction around text structure and text features to improve student comprehension of text (e.g., Englert & Thomas, 1987; Gajria et al., 2007; Faggella-Luby, Schumaker, & Deshler, 2009). Teaching students the strategy of inferencing is supported by the work of Marzano (2010), Ozgungor and Guthrie (2004), and others. Inferencing is the act of drawing a conclusion

based upon information that is implied or unknown. As student progress through the grades and text becomes increasingly complex, students often have to rely on inferencing to gain all of the important information relayed by a text. Students need to coordinate knowledge from both what is implied in the text as well as what is overtly stated in a text. As stated earlier, the act of finding main ideas within text is important for helping students construct knowledge when reading text. In addition to helping students remember the most important information from a text, this type of instruction can also help students focus upon other important aspects of text, such as the author's message or purpose. Researchers such as Winograd (1984) and Garner (1985) provide evidence to support the inclusion of instruction regarding finding main ideas important to literacy instruction.

While all of these strategies have demonstrated positive effects on improving students' comprehension of text, research has demonstrated that teachers must exercise caution in providing instruction in these strategies; while it is effective to teach several strategies to students at one time, one must avoid teaching too many strategies at once (NICHD, 2000). Packaged strategies are often a good solution to this problem in that they bundle effective strategies together and provide a framework for presenting the strategies to students in a structured way that reduces the problem of presenting too many strategies to students at one time.

Discussion of Reading Content. Discussion of reading content refers to teaching in which teachers engage students (either with the teacher or with others in the classroom) in discussion of the content of text. Teaching behaviors often encompassed in descriptions of instruction in the discussion of reading content include: using others'

questions and comments to build discussion, expressing opinions or taking a position, making connections across time and subjects, questioning the author, and asking authentic questions (e.g., Applebee, Nystrand, & Gamoran, 2003; Biancarosa & Snow, 2006; Guthrie et al., 2000; Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2007; Reznitskaya et al., 2001). Discussion promotes deep understanding of text (Applebee et al., 2003) by allowing students the opportunity to internalize the thinking processes they experience through such class discussions (Kamil et al., 2008). In such discussions, students are exposed to multiple examples of how meaning is constructed from text, express and defend their opinions of text, identify specific sections of text that defend their positions, listen to others express and defend their own, often unique, opinions, and more. Such exchanges of ideas, occurring on frequent bases, provide students exposure to and modeling of a variety of ways to extract meaning from text and improve comprehension.

Motivation. Motivation refers to instructional behaviors exhibited by the teacher that foster students' motivation to learn and engage in instruction. Teaching behaviors often encompassed in descriptions of motivating instruction include: focusing students on important and interesting learning goals, providing a range of activity choices, providing interesting texts at multiple reading levels, providing opportunities for student collaboration in discussion and assignments, and providing connections between learning and relevance to students' lives (e.g., Applebee et al., 2003; Biancarosa & Snow, 2006; Gambrell, 2011; Guthrie & McCann, 1997; Guthrie et al., 1999; Guthrie et al., 2000; Guthrie et al., 2007; Henderlong & Lepper, 2002; Schunk & Rice, 1992). Adolescent readers, when given the right situation, with a motivating text and appropriate support,

can be seen to read skillfully and attentively (Guthrie & Davis, 2003). Fostering motivation to read is essential to all readers, but is often viewed as especially important for struggling readers who are often disengaged from reading after years of reading failure (Guthrie & Davis, 2003). Reading achievement has been associated with the amount and breadth of reading one engages in (Cunningham & Stanovich, 1991); in turn, the amount and breadth of reading one engages in is also related to one's intrinsic motivation to read (Wigfield & Guthrie, 1997). Thus, assisting students in the development of motivational goals in relation to reading can contribute to reading achievement in the short term, and also lead to long-term improvements in reading by increasing students' desire to read for pleasure and purpose throughout their life span (Guthrie et al., 1999).

Technology. Technology refers to the use of technology as an instructional tool in the delivery of literacy instruction. Research increasingly suggests that including technology in instruction, whether as a vehicle with which to deliver instruction (such as computer programs, smart boards, digital projectors, etc.) or as a topic to facilitate discussions of literacy (such as how technology influences how we develop our literacy skills), is important to effective literacy instruction (see Biancarosa & Snow, 2006; Mastropieri, Scruggs, & Graetz, 2003; Swanson, 1999; Swanson & Deshler, 2003 for reviews). This aspect of quality instruction provides both an engaging topic and medium through which teachers can support student literacy.

Writing. Writing refers to teacher instruction in which the teacher develops student literacy skills through writing. Writing is typically a foundational element of a strong approach to literacy instruction (e.g., Biancarosa & Snow, 2004; Graham &

Hebert, 2010). Instructional activities in writing to improve comprehension of text include teaching students how to write about what they have read (e.g., personal reactions, analyses of text, answering questions about the text in writing, summarizing a text in writing), as well as teaching students the writing skills and processes needed to create text (e.g., spelling, sentence construction, text structures for writing) (Graham & Hebert, 2010). When students understand the skills and processes needed to create a written text, there is a related increase in reading skills (e.g., Graham & Hebert, 2010), as both reading and writing share common processes and understanding (Fitzgerald & Shanahan, 2000).

Rationale for the Present Study

With growing attention to the national literacy crisis, school psychologists need to be prepared to examine fidelity of implementation to Tier I ELA curriculum when conducting classroom observations. As the literature indicates, fidelity of implementation is an under-researched, but vital component to determining intervention effectiveness. While some information about implementation fidelity has surfaced in the medical and behavioral health fields, educational research still lags behind with regards to effectively measuring this construct. Furthermore, within the field of education, most of the implementation fidelity research has been conducted with preschool and elementary students, and has been focused on interventions, including educational (math, reading, etc.), health (preventing substance abuse, smoking, etc.), and behavioral (improving social skills, etc.). Little-to-no implementation research exists for students at the middle and high school level, particularly in the area of core language arts curriculums.

Research Questions and Hypotheses

The purpose of this study was to evaluate the correlation between teachers' fidelity of implementation of an English Language Arts curriculum and student outcomes on curriculum-based measures of reading achievement. This study addressed the following research questions:

1) Do individual teacher differences on measures of fidelity of implementation of the English Language Arts curriculum correlate with student achievement on curriculum-based measures of reading achievement?

2) Do specific aspects of fidelity of implementation to the English Language Arts curriculum (such as adherence vs. quality) more highly correlate with student achievement in reading?

The null hypothesis for the first question in this study is that there will be no statistically significant correlations among teacher fidelity of implementation of the ELA curriculum and student reading achievement. The null hypothesis for the second question in this study is that there will be no statistically significant differences among the relationships of the separate elements of teacher fidelity of implementation of the ELA curriculum (such as adherence, dosage, etc.) and student reading achievement. It was hypothesized that the null hypotheses will be rejected and that statistically significant relationships will be observed among teachers' fidelity to the ELA curriculum and adolescent reading achievement. In other words, it was hypothesized that students whose teachers demonstrate higher levels of fidelity of implementation of the curriculum will achieve higher reading performance on curriculum-based measures, and that quality of implementation will be more highly correlated with improved student outcomes for

adolescent readers than the other three aspects of fidelity of implementation examined in this study.

Chapter 2: Methods

Setting and participants

Setting. An urban public school system on the east coast of the United States was recruited to participate in this study. The school district has 12 schools (8 elementary, 2 middle schools, and 2 high schools), serving 8,279 students in grades preschool through 12 (according to currently available data from 2010; Connecticut Education Data and Research (CEDaR), 2013). As this study focuses on adolescent literacy, the two district middle schools were recruited for participation in this study. Both middle schools instruct students in grades 6 through 8; one school's student enrollment is listed at 711 total students, while the other middle school's enrollment is listed at 888 total students (based on most recent publically available data from 2010; CEDaR, 2013). Publically available information provides average class size for grade 7 (the only middle school grade available) and total hours of instructional time for English Language Arts (ELA) for grade 8 (again, the only middle school grade available). In this district, the average class size in grade 7 is 21.5 students, which is slightly higher than the state average of 20.6 students per class (CEDaR, 2013). The district devotes 957.5 hours of time to instruction yearly (compared to the state average of 1,000 hours (CEDaR, 2013)), with 272 of those hours devoted to ELA instruction (CEDaR, 2013).

Teachers. All teachers of the English Language Arts (ELA) curriculum at each grade level at each middle school were sought for recruitment into this study. According to the most recent publically available data (CEDaR, 2013) the teachers in this school district have 15.4 years of teaching experience, on average. There was some variability between the two schools, in that one school's average years of experience among its

teachers is 18.6 years, while the other school's average is 13.8 years; this is in comparison to the state average of 14.7 years of teaching experience (CEDaR, 2013). At one middle school, 76.6% of its teachers had a Master's degree or higher as their terminal degree, whereas the other middle school had 70.5% of its teachers with a Master's degree or higher (CEDaR, 2013). In general, this is just slightly below the state average of 82.1% of teachers having a Master's degree or higher (CEDaR, 20130). Teachers in the district are absent on average 5.5 days per school year, which is slightly lower than the state average of 8.3 days absent per school year (CEDaR, 2013). Finally, 11.7% of the staff at one middle school identify as belonging to a minority racial or ethnic population, while 12.6% of the staff identify as belonging to a minority racial or ethnic population at the other middle school (CEDaR, 2013). This is higher than the state average of 7.9% of the teaching population that identifies as a member of a racial or ethnic minority population (CEDaR, 2013).

Ultimately, 1 grade 6 teacher, 1 grade 7 teacher, and 3 grade 8 teachers were recruited into the study at one middle school (for a total of 5 teachers) and 4 grade 6 teachers were recruited into the study at the other middle school, for a total of 9 teacher participants. At the former school, all 5 teachers allowed two different classrooms of students to be observed, so that ultimately 14 different classrooms of students were recruited into the study. Thus, 6 grade 6 classrooms, 2 grade 7 classrooms, and 6 grade 8 classrooms were observed in this study.

Students. Publically available information indicates that of the students enrolled in the public schools in this district, 66.6% qualify for free or reduced lunch status (64.4% at one middle school, 70.5% at the other middle school, compared to the state

average of 34.4%; CEDaR, 20130). Within the district, 13.5% of students are identified as eligible for special education services, which is higher than the state average of 11.6% and the district reference group (DRG) average of 11.4% (CEDaR, 2013). Additionally, 62.1% of the district's students identify with a minority status (55.6% at one middle school, 63.5% at the other middle school, compared to the state average of 36.3%; CEDaR, 2013). Furthermore, 11.4% of the students (11.3% and 8.3%, respectively, at the two middle schools) are identified as English Language Learners (ELL) and 18.7% of the students (15.5% and 18.4%, respectively) live in homes with a non-English home language (CEDaR, 2010). Students at all grades levels in this study (grades 6, 7, and 8), on average, underperformed in regards to state assessment of reading proficiency. Compared to state averages, a greater percentage of students were below the target proficiency level at all three grade levels, and a smaller percentage of students collectively met or exceeded the targeted level of proficiency (see Table 2.1; CEDaR, 2013).

All students enrolled in the 14 classrooms of the 9 teachers recruited into the study were also recruited as secondary participants in this study. A total of 108 students were successfully recruited across 12 of the classrooms in the study; no students were successfully recruited from either classroom of one of the grade 8 teacher participants. Many of the students recruited into the study were missing one or more of the three CBM-R measures, resulting in a final participant count of 27 students among the 6 grade 6 classrooms, 11 students among the 2 grade 7 classrooms, and 24 students among 4 of the 6 grade 8 classrooms.

Table 2.1

Student Reading Performance as Measured by State Assessment[a]

	Below Basic	Basic	Proficient	Goal	Advanced
Grade 6 District	18.1	13.8	14.3	40.1	13.6
Grade 6 State	7.3	6.1	10.5	45.5	30.5
Grade 7 District	22.4	11.0	10.4	39.2	16.9
Grade 7 State	8.8	5.5	7.8	42.0	35.8
Grade 8 District	23.6	10.4	14.5	35.3	16.2
Grade 8 State	10.4	6.2	8.7	42.3	32.4

[a]*Note*. 2010-2011 data presented (most currently available). Percentage of students meeting requirements listed for each performance category.

Participant consent. Using a protocol approved by the Human Subjects Institutional Review Board (HSIRB), written consent to participate in this study was obtained by all teachers, students, and parents of students (who are not of the age of consent) prior to the beginning of the study.

English/Language Arts curriculum description

Every teacher of the ELA curriculum for grades 6-8 are provided with a binder entitled, "Middle Level Language Arts Curriculum Resource Binder Grades 6-8" and access to an electronic file-sharing database in which many electronic resources are housed. The intended purpose of the curriculum is to provide teachers with written guidelines for providing high-quality ELA instruction for students in grades 6-8. The

binder with which the teachers are provided is divided into many sections, including sections for Standards, Reading, Writing, Assessments, and Interventions.

Standards. In this section, teachers can find the curricular standards upon which the curriculum is based. In the first segment of this section of the binder, the state-adopted Common Core State Standards (CCSS; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), including the Introduction to the standards, the K-5 ELA standards, the 6-12 ELA standards, and the 6-12 Literacy in History/Social Studies, Science, and Technical Subjects standards, are provided. Next, the state's ELA curriculum standards (adopted in 2008), including the Introduction to the standards, the specific grade 6-8 ELA standards, and the framework standards are provided. Then curriculum mapping is provided in which the CCSS standards, matched to a state grade-level expectation, are listed, with space provided for teachers to indicate when and where they taught the standard and what materials they used to teach the standard. There are separate sections for Reading and Writing provided in the curriculum mapping.

Reading. This section of the ELA curriculum binder is divided into ten headings: Non-fiction Units, Novels, Open-Ended Response Items, PowerPoint Presentations, Reading Workshop, Accelerated Reader, Assessment related to the Accelerated Reader, Summer Reading, Vocabulary, and a Walkthrough Checklist.

Non-fiction Units. For each grade level 6-8, teachers are provided with 2-3 selections. For each selection, there is an Overview page that lists the Essential Questions, Big Ideas, and Unwrapped Grade Level Expectations (with prioritized expectations listed in bold type) for each selection. Each Overview page is then followed

by a Mini-Unit "At A Glance" that provides the "Unwrapped" Grade Level Expectations, Vocabulary, Open-ended Items, and Critical Thinking Activities for each lesson from the selection.

Novels. At each grade level 6-8, a chart is provided that lists texts in the following categories: Biography, Classical Fiction, Contemporary Fiction, Historical Fiction, and Non-Fiction. The categories are further subdivided into "Accelerated" and "Academic" categories. At grade 6, no novels are listed in the Biography category, and grade 7 does not list any novels in the Contemporary Fiction category; grade 8 provides novels in all of the above-mentioned categories.

Open-Ended Response Items. This section provides a scoring rubric that describes 0, 1, and 2-point scoring criteria to use when evaluating open-ended responses to reading selections.

PowerPoint Presentations. This section simply provides a note to teachers that these materials can be found in the electronic file-sharing database.

Reading Workshop. This section simply provides a note to teachers that these materials can be found in the electronic file-sharing database.

Accelerated Reader. Here teachers can find a description of the program that the school district has adopted to guide teachers and students in the selection of independent reading materials. Included along with the description of the program are instructions for how to use the computer interface that allows teachers to enter student data; by entering data, the computer is able to match students to appropriately leveled independent reading texts. Finally, information to assist teachers in understanding and interpreting the information generated by the computer program is also provided.

Assessment related to the Accelerated Reader. This section describes the assessment component of the above-described district-adopted reading program. The description provided indicates that students should be evaluated 3-4 times per school year so that "just right" texts can be assigned to students for independent reading. The assessment provides scores related to zones of proximal development that are linked to levels of texts used in the program.

Summer Reading. While this section does not include the actual list of grade-appropriate summer reading texts, it does provide teachers with sample letters that can be sent home to parents informing them of the summer reading program, blank reading logs that can be distributed to students to help them track their summer reading, and blank certificates that teachers can customize and distribute to students to celebrate their summer reading accomplishments.

Vocabulary. A list of all new vocabulary words that are to be taught during the course of the school year is provided in this section. Each grade level has 20 separate lists of vocabulary words; each list is divided into 4 sections: Integrated Language Arts, Math, Science, and Social Studies. Each of the 4 sections of each list include 5 words specific to that subject area.

Walkthrough Checklist. Here teachers are provided with a copy of the checklist that supervising school administrators will use when performing evaluations of their teaching.

Writing. This section of the ELA curriculum binder is divided into four headings: Composing/Revising Checklists, John Collins, PowerPoint Presentations, and Teacher-Created Writing Resources.

Composing/Revising Checklists. In this section, teachers are provided with a map that details which elements of Composing/Revising section of the statewide assessment are measured in each grade level. Each element of the Composing/Revising element of the statewide test is listed, and an "X" is placed in corresponding grade-level columns to indicate if that element is measured in the statewide assessment at that grade level.

John Collins. Here teachers can find a Middle Level Writing Plan from the district-adopted Collins Writing Program (Collins, 2007). Information is provided so that teachers have an understanding of how to use the Collins Writing Program (Collins, 2007) in grades 6-8. Teachers are also provided with a copy of the current school year's Academic Writing Requirements outlined by the school district per grade level, per subject area.

PowerPoint Presentations. This section simply provides a note to teachers that these materials can be found in the electronic file-sharing database.

Teacher-Created Writing Resources. This section simply provides a note to teachers that these materials can be found in the electronic file-sharing database.

Assessments. This section of the ELA curriculum binder is divided into four headings: the statewide assessment, Critical Reasoning Exercises (and Degrees of Reading Power), District Writing Assessment, and the district-wide assessment.

Statewide assessment. In this section teachers are provided with information related to each separate assessment administered by the state in the yearly assessment of student academic achievement given to meet the requirements of the No Child Left

Behind legislation (No Child Left Behind [NCLB], 2003). Thus, test information specific to reading, writing, math, science, and other measured areas are provided.

Critical Reasoning Exercises (and Degrees of Reading Power). Here teachers can find an Introduction and PowerPoint presentation that provides information related to critical reasoning exercises that the school district requires teachers to use with students, as well as information related to the degrees of reading power (DRP) system for assessing student reading levels.

District Writing Assessment. The district administers an assessment of each student's writing performance three times throughout the school year. The corresponding Fall, Winter, and Spring writing prompts are provided for each grade level in this section.

District-wide assessment. The school district has created their own measures of student academic performance that are *in addition to* the statewide assessments administered throughout the school year. In this section, teachers are again provided with the Open-Ended Questions Scoring Rubric (also described in the earlier Reading section of the binder), as well as an Open-Ended Item Response Sheet.

Interventions. Interventions that the district uses to assist students in the development of their ELA skills are provided here. Some of the interventions listed include Read 180, Read About, System 44, and Write to Learn. It is noted here that all of the materials for these interventions can be found in the electronic file-sharing database.

As noted above, teachers are also provided access to an electronic file-sharing database. In addition to *all* of the materials described above located in the binder, the file-sharing database also includes the following resources:

- Appendices A-C of the Common Core State Standards

- Additional lesson plans, activities, and resource materials
- Assessment materials (such as Open-Ended question items, Anchor sets, and rubrics)
- Non-fiction mini-units
- Pacing guides for the Common Core State Standards in writing (the writing curriculum broken down into two-month pacing guides, as well as instructions for how to transition the pacing guides to the reading curriculum)
- Reading PowerPoint Presentations
- Reading Workshop (including materials to support set-up and use of reading workshop with the reading curriculum (mini-lessons))
- Reports (of student assessment and performance data)
- Sample multiple-choice items for the Editing/Revising elements of writing (similar to the statewide assessment of writing in the area of Composing/Revising)
- Teacher-Created Writing Resources to share
- Writing materials (including assignments, Anchor sets, Holistic scoring rubrics, student writing samples)
- Writing PowerPoint Presentations

Materials needed to implement the curriculum. In order to implement the above-described school district's ELA curriculum, teachers need access to the binder containing the written ELA curriculum, a computer and internet capability to access the file-sharing database that supports the curriculum, all of the texts used to deliver the curriculum, and

all of the district-adopted programs (e.g., John Collins, Read 180) that are used to implement and support the curriculum.

Training provided to teachers to orient to the curriculum. Prior to implementing the ELA curriculum for adolescents for the first time, every teacher in the school district receives the binder of curriculum materials and access to the file-sharing database to orient himself or herself to the curriculum. New teachers are also provided training by the Supervisor of Language Arts (grades 6-12) and the school-based reading specialist in best practices for delivering the curriculum. Finally, the school-based reading specialist schedules himself or herself into the new ELA teacher's classroom for the first term of the school year in a coaching and co-teaching model to provide direct support in delivering the ELA curriculum.

Measures

Student outcomes. The school district has a department of Research and Evaluation that is devoted to assessing and improving the educational progress of all students. The specialists in the Research and Evaluation Office (REO) were tasked with the assignment of creating assessments that both fulfilled the yearly assessment requirements of NCLB legislation and provided teachers with timely and useful information to guide their instructional practices (Larson & Kelleher, 2009). As such, the specialists in the REO decided to develop diagnostic assessments that were similar to Cognitive Diagnostic Assessments (CDAs). CDA is an approach to assessment that provides "fine-grained" formative diagnostic feedback about learners' mastery of skills assessed (Jang, 2008). By employing a CDA, one can determine the processes and strategies test takers employ to solve the problems with which they are presented (Jang,

2008). By understanding the skills used to successfully and unsuccessfully answer test items, teachers can inform their instructional practices by targeting areas of weakness, as evidenced by inappropriate strategy use in solving test items incorrectly. The use of CDA helps teachers understand not only *what* students don't know in a particular curricular area, but also informs them of the incorrect strategies students employ to arrive at incorrect answers (Jang, 2008). Using this theoretical approach to test development, the specialists in the school district's REO developed district-wide Error Identification Assessments (EIa) to provide meaningful instructional information to teachers while monitoring the progress of students in attaining district-set benchmarks for educational achievement.

District-developed EIa. When applying the CDA model to the assessment of reading, the specialists in the REO developed an online diagnostic EIa for all students in grades 3-10. The EIa is administered 3-4 times per year to every student in grades 3-10 to assess their progress in the development of reading skills. As originally developed, the EIa in reading is comprised of 46 original reading comprehension passages, with 454 reading comprehension questions (Larson & Kelleher, 2009). For each reading comprehension item, the test developers provide "one correct, or "most correct" answer, with four foils, each designed to mimic a typical student misconception" (Larson & Kelleher, 2009). The EIa reading assessment is printed in an "attractive format for both teachers and students" by the district's REO using Microsoft Publisher Software (Larson & Kelleher, 2009). Although it is an online assessment, students are provided with test booklets, in which they may write or underline information; however, students must "record their answers on a customized scannable answer document (bubble sheet) that is

'pre-slugged' (utilizing Design Expert Software) with each student's name, identification number, school and teacher" (Larson & Kelleher, 2009). To ensure that the EIa truly assesses students' skills and reduce confounding variables such as time, students are provided with flexible time limits, and are able to take out-of-level tests, as appropriate. Once the student completes the assessment, the student's answer documents are "scanned (utilizing Scan Tools Plus Software) and cleaned ("double bubbles" due to poor erasers are identified using SPSS programming) and hand corrected in order to report what each student was thinking during testing" (Larson & Kelleher, 2009). Once scanned and analyzed, the SPSS programming produces a "series of reports, listing error and pattern of errors by students within classrooms, as well as Individual Error Reports for each student in reading comprehension" (Larson & Kelleher, 2009). Since the EIa were developed to inform instruction and provide immediate feedback on specific student errors, the REO makes every effort to return error reports to teachers within 3 to 5 days.

According to the district's Research and Evaluation Office, the district's reading EIa was "constructed and revised utilizing traditional approaches to test development" (Larson & Kelleher, 2009). Research from the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (1999); Kolen and Brennan (2004); and, Feuer, Holland, Green, Bertenthal, and Hemphill (1999) (all as cited in Larson & Kelleher, 2009) indicates that EIa statistically links to the scale scores of the state mandated reading tests. According to Larson & Kelleher (2009), "[t]his calibration also provided an EIa scale score and NCLB performance labels (Below Basic, Basic, Proficient, Goal and Advanced) for each student." Providing these labels can be a convenient way of

communicating student performance, as teachers are already familiar with these performance labels from the state mandated testing. The specialists in the district REO found that their EIa correlated well with state assessments (R range .85 to .93 in reading), and as such, could be used as "independent variables in multiple regression equations to predict future performance on state required tests" (Larson & Kelleher, 2009). Due to the reliability of the EIa to serve a predictive function for future reading performance, they provide a lens through which teachers can identify students who are at-risk for failing to meet established district reading performance benchmarks, and can provide targeted intervention to assist students in improving their reading performance. The error analysis element of the EIa allows teachers to more accurately select intervention tools to target areas of weakness in students' development of reading skills, as well as inform their general teaching practices.

Teacher fidelity of curriculum implementation. All ELA teachers have been provided with a written ELA curriculum that outlines the expectations for student ELA achievement for the school year. To document the fidelity of teachers' implementation of the ELA curriculum, teachers will be observed while implementing ELA lessons in their classrooms. Observation was selected as the method for collecting data regarding teacher fidelity of implementation of the ELA curriculum because direct observation is considered a first-level approximation of student learning (Kennedy, 1999) and because it is less prone to self-report bias (Ransford, Greenberg, Domitrovich, Small, & Jacobson, 2009).

The instrument and coding for the instrument was developed, in part, based upon a similar instrument used in another research study (Cornett, 2010). Keeping in mind

what the research suggests to be the five critical components of fidelity of implementation (adherence, dosage, quality, differentiation, and student responsiveness), the adopted instrument was developed to document four of these five components (adherence, dosage, participant responsiveness, and quality). The fifth element of fidelity of implementation, program differentiation, will not be measured at this time, as this study will not look to examine the essential elements of an ELA curriculum that need to be implemented for successful outcomes.

Adherence. Based upon a review of the fidelity of implementation literature, adherence is conceptually defined as how closely the steps of a program are followed. Considering the ELA curriculum to be a type of "program", adherence was then operationally defined as documentation of the number of lessons ("steps") delivered according to the scope and sequence of the ELA curriculum.

Dosage. Based upon a review of the fidelity of implementation literature, dosage is conceptually defined as the amount of the program delivered. Again, considering the ELA curriculum to be a type of "program", dosage was then operationally defined as the number of minutes of a given ELA class period devoted to delivering lesson content.

Participant responsiveness. Based upon a review of the fidelity of implementation literature, participant responsiveness is conceptually defined as the degree participants respond to, or are engaged by, a program. With this conceptual definition, participant responsiveness was then operationally defined as student engagement in literacy instruction.

Quality. Based upon a review of the fidelity of implementation literature, quality of implementation is conceptually defined as how well the program is delivered. Using

this conceptual definition, quality of implementation was then operationally defined as the effective, research-supported instructional procedures teachers use to deliver curriculum to adolescent readers. To understand what these effective, research-supported procedures are, existing literature was reviewed. The literature search was conducted using the following online databases: Academic Search Premier, Education Research Complete, ERIC, MasterFILE Premier, PsycARTICLES, PsycINFO, and University of Connecticut eBook and Professional Development Collections. To identify scholarly articles and book chapters relevant to this literature, the following keyword search terms were used: adolescent, effective instruction, effective teaching, English, high school, instructional practice, literacy, literacy instruction, quality instruction, quality teaching, reading, secondary, and student achievement. From the corpus of literature returned, seminal articles were identified to perform ancestral searches.

This literature was examined to identify quality instructional behaviors to support the delivery of the ELA curriculum for adolescent readers. As such, six categories of quality implementation of instruction for adolescent readers were identified: time spent in transition, engagement (student and teacher), teaching function, management of student behavior, assessment, and quality instruction components (such as the use of technology, strategy instruction, etc.). Once these categories were identified, lists of activities that described the categories were generated from the literature to begin to create an observational tool in which the directly observable behaviors of the categories of interest were labeled. In all, 22 activities were ultimately identified to describe the above categories of quality of implementation. While using these 22 behaviors to create the observational tool, it was discovered that the original six categories of quality

implementation of instruction could be collapsed into five broad categories for the purpose of this study, namely: teaching function (which included formal assessment), student engagement, teacher non-instructional behavior (includes teacher engagement, time in transition, and management of student behavior), feedback (which included elements of informal assessment), and quality instruction components. These resultant five categories were sub-divided into the specific activities that comprised the over-arching category; the categories and accompanying activities were used to create the observational tool used for this study (Appendix B). Below are brief descriptions of each category; operational definitions and examples of teaching behaviors that demonstrate the element of quality instruction used as decision criteria for observers using the observation tool can be found in Appendix C.

Teaching function. Teaching function (named TF in data analyses) consisted of eight sub-categories: presentation of an advance organizer, review, presentation of a demonstration, presentation of a model (modeling), guided practice, peer practice, independent practice, and formal assessment.

Teacher non-instructional behavior. Teacher non-instructional behavior (named NI in data analyses) consisted of four sub-categories: time in transition, telling anecdotes, reading aloud to students/monitoring student work (without interacting or providing feedback), and off-task behavior (such as checking email, talking on the telephone, etc.).

Feedback. Feedback (named FEED in data analyses) consisted of three sub-categories: questioning students to probe understanding, acknowledging student responses (without providing elaborative feedback), and elaborative feedback.

Quality instruction components. Quality instruction components (named QI in data analyses) consisted of seven sub-categories: content knowledge instruction, word reading/vocabulary instruction, comprehension strategy instruction, discussion of reading content, motivation, writing, and technology.

Student Engagement. From an observational standpoint, engagement is difficult to observe (e.g., Fredricks, Blumenfeld, & Paris, 2004); that is, it is difficult to determine a student's engagement in instruction with certainty through observational methods. Thus, in order to facilitate data collection of this important aspect of quality instruction, this category needed to be defined in its inverse. Since one can fairly reliably observe behaviors that represent disengagement in instruction, engagement was defined as the *absence of* disengaged behaviors.

Procedures

Teacher fidelity of curriculum implementation. The literature regarding procedures for observing teachers delivering instruction remains uncertain as to what a standard number of observations would be to ensure reliability of observed behaviors (Domitrovich et al., 2008). Currently, three to five observations seem to be an acceptable minimum, but up to ten observations may be warranted in certain cases (Rowan, 2005; Erlich & Shavelson, 1978; Shavelson & Dempsey, 1976). This research suggests that it may be appropriate to increase the number of observations from the recommended minimum in cases where the observed behavior is highly variable, such as teaching behaviors across an enacted curriculum. Additionally, it is unclear what the impact of the timing of the observation (in relation to time-of-day, relationship to location within the

scope and sequence of the curriculum, etc.) or the length of the observation contributes to the reliability of observed behaviors (Domitrovich et al., 2008).

Two independent observers (including this student researcher) conducted the observations of teachers implementing the ELA curriculum. Both observers were trained in the coding scheme of the observational tool and partial interval recording (PIR) and momentary time sampling (MTS) procedures. First, both observers read and discussed the operational definitions and examples of each category and sub-category of quality instruction defined in this study, as well as the operational definitions of adherence, dosage, and student engagement. Then the observers practiced using the observational tool with publically available video recordings of ELA instruction for adolescent readers of teachers and students not involved in this study. Once the observers were in 90% agreement in the four elements of teacher fidelity of implementation of curriculum defined in this study, data collection was scheduled. Procedures specific to data collection for each element of fidelity of implementation are described below.

Adherence. At the top of the observational tool, a space is provided for the observer to record the date and topic of the lesson the teacher implements during the observation. This information was used to match the topic of the lesson to the scope and sequence of the enacted curriculum to determine the observed teacher's adherence. The topic of the lesson needed to match the alignment of the topic to the scope and sequence; that is, the topic being taught must align with the pacing of the curriculum described in the ELA scope and sequence. Adherence was coded as a dichotomous variable, in that the teacher was either in adherence to the curriculum or was not in adherence to the curriculum. This adherence variable was documented in the data analysis as a

percentage; thus, as this is a dichotomous variable, for each observation the teacher was either in 100% or 0% adherence to the curriculum.

Dosage. At the top of the observational tool, space is provided for the observer to record the start time and end time of the lesson. Since this study is interested in documenting time in transition within the observed classrooms, the observers will indicate the start time as the scheduled time that the lesson is to begin (based upon the school's schedule on the day of the observation) and indicate the end time as the scheduled time that the lesson is to end (again, based upon the school's schedule on the day of the observation). Observational intervals recorded as time spent in transition, teacher not engaged in instruction, or other teacher non-instruction activities, will be deducted from the time devoted to the class period. In an attempt to capture the discrete ways in which a teacher could be considered engaged in non-instructional behavior, four sub-categories were created. One category, named Off-Task, captured behaviors such as: talking with another adult (about something other than instruction (conversing with a co-teacher to troubleshoot instruction shall NOT be considered unengaged)), talking on the telephone, grading papers, and other non-instructional behaviors (see code book in Appendix C for full definition). Another category, named Anecdote, captured behaviors such as engaging in discussion with students not grounded in text or managing student behavior. A third category, named Transition, captured observational intervals in which the class was transitioning from one activity to another. This included transitioning into the classroom and getting settled for instruction, transitioning from one learning activity to another during the class period, and other transitional times. Finally, a teacher could be deemed as not engaging in instruction by monitoring students without providing

feedback. Examples of this type of behavior include checking homework completion, passing out papers, taking attendance, or reading aloud to students. Despite evidence to suggest that teachers reading aloud to students may improve comprehension, increase vocabulary and listening skills, and increase student motivation to read (Ariail & Albright, 2006; Ivey, 2003; Routman, 1991), research is scarce to document that it has any significant impact on reading achievement for adolescents (Cantrell et al., 2014). While an argument could be made that this behavior could be considered a type of modeling, and as such coded among the teaching functions described earlier, since this practice is not currently supported by research, it was included in the teacher non-instructional behavior (NI) category for the purposes of this study. This sub-category was labeled Monitoring/Reading Aloud on the observational tool.

Data collection for dosage and quality of implementation (see below) will be conducted in real-time using PIR, beginning and ending according to the scheduled class period of the observed class on the day of the observation. The PIR procedure will be conducted in real-time, beginning and ending according to the scheduled class period of the observed class on the day of the observation. PIR is a method for observing and recording behavior that involves segmenting large blocks of time, such as a class period, into small time intervals (e.g., 30 seconds) and indicating whether target behavior(s) occur at any time during the observed interval (Cooper, Heron, & Heward, 2007). PIR was chosen as the observational technique for the dosage and quality of implementation variables in this study over other common observational methods used in education, such as MTS, because PIR allows for the observation of multiple behaviors concurrently (Cooper, Heron, & Heward, 2007). Since the dosage and quality of implementation

variables had many behaviors that were identified as representing the over-arching domains, there were multiple activities that needed to be able to be observed concurrently. In all, there were 22 separate behaviors that defined the dosage and quality of implementation variables. As the observers needed to determine which of the 22 behaviors occurred within every 30-second interval, it was necessary to use PIR rather than MTS methods for these behaviors. It is not practical, or possible, to note whether 22 behaviors are or are not occurring concurrently at the end of a 30-second interval. As such, using PIR and noting whether any of the 22 behaviors occurred at any time during the 30-second interval, was a far more practical and accurate approach to the collection of this data.

Beneath the space provided to indicate elements of adherence and start and end time of the observation, the observational tool contains a grid which will be used to record the presence of non-instructional behaviors or transition time at each designated time interval. The observer(s) will record which, if any, of the non-instructional behaviors are present at any time during each observational segment; as such, for each observational segment, it is possible to record one, more-than-one, or no non-instructional behaviors. Likewise, the observers will also record whether the students/class was engaged in transitioning to or from an activity at any time during each observational segment. Thus, during the course of the entire observational period, observational segments in which the teacher is engaged in any of the noted non-instructional behaviors or in which the class or students are in transition will be subtracted from the total instructional time. Dosage will then be calculated as the

percentage of time devoted to instruction, minus non-instructional and transition segments, within the observational period.

Quality. Using the grid described above, the observers will also record the presence of quality indicators of adolescent literacy instruction at each designated time interval. The observer(s) will record which, if any, of the quality indicators are present during each observational segment; that is, for each observational segment, it is possible to record one, more-than-one, or no quality adolescent literacy instruction indicators. The quality of instruction variable will be entered into the data analysis as the percentage of PIR intervals where at least one of the indicated quality adolescent literacy instruction elements was recorded as present.

Participant responsiveness. Data collection for participant responsiveness (student engagement) will be collected using the MTS observational recording procedure. The MTS procedure will be conducted in real-time, beginning and ending according to the scheduled class period of the observed class on the day of the observation. MTS is a method for observing and recording behavior that involves segmenting large blocks of time, such as a class period, into small time intervals (e.g., 30 seconds) and indicating whether target behavior(s) occur at the moment that each time interval ends (Cooper, Heron, & Heward, 2007). MTS was chosen as the observational technique for the participant responsiveness variable in this study over other common observational methods used in education, such as PIR, because MTS is less likely to over-estimate the occurrence of target behaviors than other techniques. In fact, research indicates that when using short intervals (30 seconds or less), MTS provides a reasonably accurate estimate of behavior (Gardenier, MacDonald, & Green, 2004; Murphy & Goodall, 1980;

Powell, Martindale, Kulp, Martindale, & Bauman, 1977). Additionally, as the student engagement variable was a single variable without sub-component behaviors, it was practical to observe whether this single behavior occurred at the end of each interval (as opposed to the dosage and quality of implementation variables, described above, that had multiple sub-component behaviors that were not able to be observed concurrently).

Nested within the grid that captures the presence of indicators of quality instruction of adolescent literacy, the observational tool also contains space to record whether the designated student is observed to be disengaged or not at the end of each MTS interval (see previous description of engagement category). In order to ensure that a sufficient number of observations were conducted per student observed, three students were randomly selected to be observed from the students recruited into the study in each classroom. The same three students were observed on all three observation occasions in each classroom. When the trained observers conducted their observations of the classrooms, they first ensured that they were positioned such that he/she could observe both the teacher and the three target students in the classroom. The observer(s) began by assigning each of the three students a number from 1 through 3, and observing the three students in order (student 1 at the end of the first interval, observing student 2 at the end of the second interval, and student 3 at the end of the third interval) and then repeating the sequence, in the same order, until the end of the designated observational period.

At each designated MTS interval, the observer(s) will record if the designated student is *disengaged* from instruction. The participant responsiveness variable will be entered into the data analysis as the percentage of MTS intervals where students demonstrated disengaged behavior subtracted from 100%. This will yield an estimate of

the total percentage of time students were likely engaged in instruction (evidenced by the absence of disengaged behavior).

Data collection. Due to staffing constraints (the student investigator and an additional student researcher were available to collect data for this study), each teacher recruited into this study was observed on three separate occasions to observe the fidelity with which they implemented the designated ELA curriculum. To assist with ensuring a high degree of inter-rater reliability across the two observers over the course of the study, 10% of the teachers recruited into the study (n=2) were randomly selected to be observed by both trained observers on all three observation occasions. The remaining teachers (n=12) were observed twice by each observer, resulting in overlap on one occasion for each teacher. This overlap will provide an opportunity to calculate inter-rater agreement on one-third of the observations in the study. To ensure a systematic, counter-balanced approach to the overlap conditions, the teachers were randomly assigned to one of four overlap observation conditions: ALL (overlap on all three observations), 1st (overlap on the first observation), 2nd (overlap on the second observation), or 3rd (overlap on the third observation). This yielded 42 observations of the 14 teachers in this study, with each observer conducting 30 observations.

Inter-Observer Reliability. As noted above, to determine the reliability of ratings completed between the two independent observers, 10% of the teachers recruited into the study (n=2) were randomly selected to be observed by both trained observers on all three observation occasions. The remaining teachers (n=12) were observed twice by each observer, resulting in overlap on one occasion for each teacher. For all observations in which each observer was present and independently observed the classroom instruction

(n=18, 43% of the total 42 observations), inter-observer percent reliability agreement could be calculated. Inter-observer percent reliability agreement was calculated using the following formula: Percent Reliability = (Number of Agreements / Number of Agreements + Disagreements) X 100. Inter-observer agreement across all intervals was 98.6 percent reliability, suggesting adequate agreement for the research purposes of this study.

Overall fidelity of implementation of ELA curriculum. From the collective data obtained regarding the elements of adherence, dosage, quality, and participant responsiveness, it will be possible to calculate an overall fidelity of implementation score for each teacher. Each of the four elements was equally weighted, contributing 25% to the overall fidelity of implementation score of each teacher. This overall fidelity of implementation score allowed for comparison between teachers who demonstrate differences in their overall fidelity of implementation.

Student outcomes. In this study, student reading achievement was assessed by school district-developed Error Identification Assessments (EIa; see above). EIa collected in May of the previous school year (May 2012) served as a baseline for student reading achievement. EIa administered at the start of the school year (September 2012) and just prior to data collection (December 2012) were also collected to assess student reading achievement in relation to the instruction observed during the scheduled teacher observations. These EIa measures were then used to determine if relationships existed between the independent teacher variables (teaching function, teacher non-instructional behavior, feedback, and quality instruction components) and student engagement in the ensuing data analyses.

Data Analysis

First, the student EIa data were analyzed to ascertain that there was, in fact, variability among the students' reading performance. To determine the existence of this variability, a repeated measures ANOVA was conducted, examining variability in the students' May 2012, September 2012, and December 2012 EIa scores. Once statistically significant variability was established, the EIa measure was compared to school, student grade-level, and classroom level variables to identify at which levels (school, grade, or classroom) the variability existed. Since these data for this study is hierarchically nested (students are nested within teachers' classrooms, which are nested within grade levels, which are nested within schools), determining the level (school, grade, and/or classroom) at which the variability existed was critical to employing the most efficient and appropriate analysis of the independent teacher variables. A repeated measures ANOVA was again performed to examine the relationship between the students' EIa data and the SCHOOL variable (coded 0,1 to represent the two different schools in the study), the GRADE variable (also simple coded to represent the three grade levels in the study), and the CLASS variable (which was also simple coded and represented the 12 classrooms in the study that had student EIa data to compare to these variables).

Once the general variability in the student EIa data were identified, repeated measures ANOVAs were performed to compare the student EIa data to the independent teacher variables. These analyses were conducted to determine if a statistically significant relationship existed between the student EIa data and the observed independent teacher variables. The first analysis compared the student EIa data to the global measure of fidelity of implementation, TOTFID. Subsequent analyses compared

the student EIa data to the four constituent elements of fidelity of implementation examined in this study: adherence, dosage, quality, and student engagement. Since the quality variable also had many constituent elements that comprised the overall quality variable (TOTQUAL), this element of fidelity of implementation was also further explored to determine if statistically significant relationships existed between the elements of TOTQUAL (e.g., teaching function) and the student EIa data, again using repeated measures ANOVAs.

As noted, the student EIa data collected for this study are hierarchically nested; students are nested within teachers' classrooms, which are nested within grades, which are nested within schools. Thus, hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) was also used to analyze data collected in this study. Variables found to have statistically significant relationships in the above-described analyses were then brought into HLM to explore whether or not the independent teacher variables could demonstrate a statistically significant proportion of the variance in the outcome measure.

Chapter 3: Results

The purpose of this study was to evaluate the correlation between teachers' fidelity of implementation of an English Language Arts curriculum and student outcomes on curriculum-based measures of reading achievement. This study addressed the following research questions:

1) Do individual teacher differences on measures of fidelity of implementation of the English Language Arts curriculum correlate with student achievement on curriculum-based measures of reading achievement?
2) Do specific aspects of fidelity of implementation to the English Language Arts curriculum (such as adherence vs. quality) more highly correlate with student achievement in reading?

To answer these questions, curriculum-based measures of student reading achievement developed by the school district (Error Identification Assessments (EIa)) were collected from May 2012, September 2012, and December 2012 of students successfully recruited into the study from the teacher participants' classrooms (see previous chapter for further details). An observational tool was developed (see Appendix C) to collect data regarding teacher fidelity to the ELA curriculum; as such, data regarding the adherence, dosage, quality of instruction, and student engagement in each observation was collected. These measures of fidelity of implementation were collected through direct observations of classroom English Language Arts (ELA) instruction. As described in chapter 3, a total fidelity of implementation score (TOTFID) was calculated for each observation, and across observations for each teacher. Teachers' instructional behavior (see Appendix C for detailed descriptions) was collected in the following

categories: teaching function, non-instructional behavior, feedback, and quality instruction indicators. Each of these categories was comprised of several behaviors that were individually recorded as present or absent during each sampling interval of each observation. Randomly selected students (from the sample of recruited students in each classroom) were also observed at specified sampling intervals to record the presence of off-task behavior. These data were then used to examine the correlations between the collected variables and reading achievement. Next, the variables that seemed to be more related to reading achievement were examined in more detail by exploring the relationship between these variables and student engagement.

First, the data on students' reading achievement is presented, with the data depicting teacher fidelity next, followed by correlations between the teacher fidelity and student reading achievement, and finally exploratory analyses related to the relationship between two aspects of fidelity: teacher quality and student engagement. Figures 3.1-3.5 show graphs of the dependent variable compared to important independent variables. Table 3.1 presents the means, standard deviations, and ranges of the dependent EIa variable, whilst Tables 3.2-3.6 present the means and standard deviations for each independent variable for each analysis. Tables 3.7-3.9 display results from analyses performed with hierarchical linear modeling (HLM).

Table 3.1 lists the descriptive statistics for the dependent variable (Error Identification Assessment (EIa) scores) at the three grade levels (6, 7, and 8, respectively) for each time point (May, September, and December 2012) at which the students' reading achievement was measured. To ensure variability within students' reading skills, a repeated measures ANOVA was conducted, examining variability in the students' May

2012, September 2012, and December 2012 EIa scores. The EIa measure was compared to student classroom assignment to investigate the existence of differences in reading scores on these variables. Classroom was entered as a between subjects variable into a model which demonstrated a statistically significant quadratic main effect for the EIa measure (F= 16.467, p<0.001, df =1, 52), indicating a statistically significant curvilinear effect for the EIa measure across time (see Figure 3.1). For these data, the reading scores were higher at the initial and final time point than at the midpoint in each classroom.

In the model where classroom was entered as a between subjects variable, results indicated a statistically significant main effect for classroom assignment (F=2.913, p=0.009, df=8, 52), demonstrating a statistically significant effect for the EIa measure across classrooms (see Figure 3.2). These data indicate that there is sufficient variability in reading performance among the classrooms in this study to warrant further investigation.

Table 3.2 lists descriptive statistics for the teacher fidelity variables across all teacher participants across all grade levels and across all observations. Tables 3.3-3.5 list descriptive statistics for the total fidelity and four component elements of teacher fidelity at the three grade levels (6, 7, and 8, respectively) for each teacher participant in the study. Total fidelity was calculated by giving equal weight to a teacher's adherence, dosage, quality of instruction, and student engagement variables. As described in the previous chapter, each of these variables contributed 25% to the total fidelity score calculated for each teacher. Examination of the relationship between teachers' total fidelity score and student EIa performance using repeated measures ANOVA failed to yield statistically significant results (reading * total fidelity linear result F=0.000, p=.988,

df=1, 59; reading * total fidelity quadratic result *F*=0.399, *p*=.530, *df*=1, 59). Thus, a teacher's overall fidelity of implementation was not found to have a statistically significant relationship to student reading achievement.

Table 3.1

Means, Standard Deviations, and Ranges for Student EIa Scores[a]

	Overall Sample (*n*=61)			
	M	*SD*	*Min*	*Max*
May 2012 EIa	229.27	33.67	166	298
September 2012 EIa	210.81	25.71	170	275
December 2012 EIa	223.88	28.53	164	296

By Grade Level

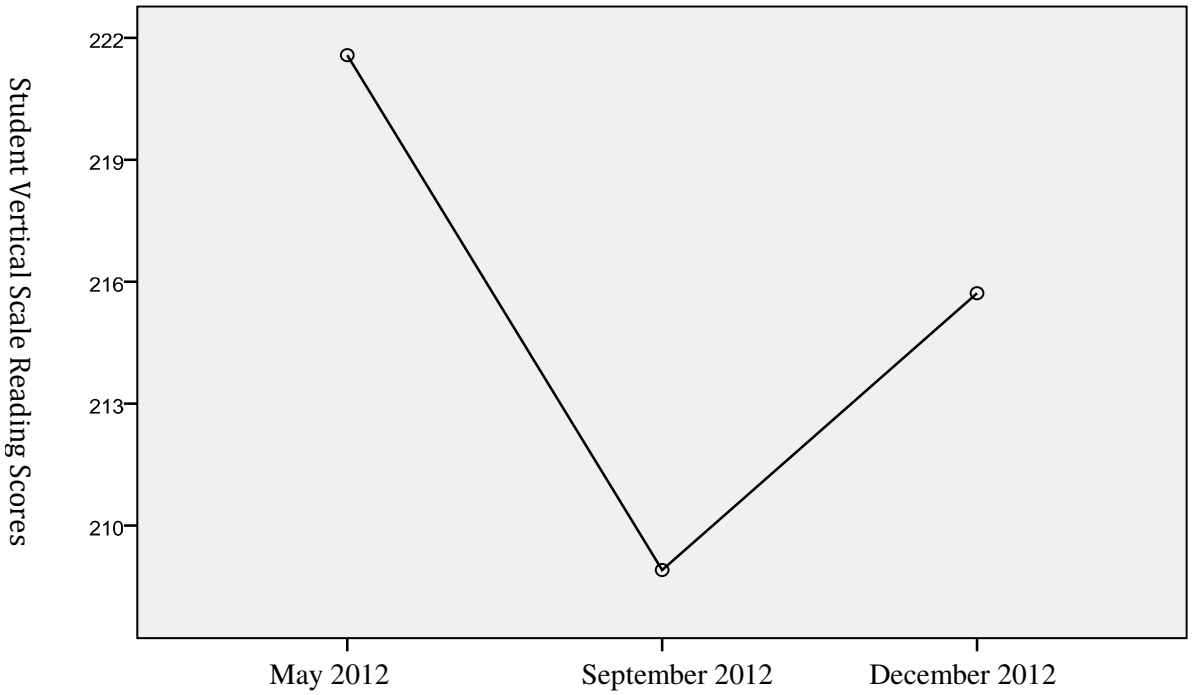
	Grade 6 (*n*=27)				Grade 7 (*n*=11)			
	M	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*
May 2012 EIa	223.74	26.99	188	278	219.73	27.45	188	296
September 2012 EIa	211.59	26.76	180	275	208.50	15.50	183	228
December 2012 EIa	228.07	28.76	186	288	205.82	17.13	189	240

By Grade Level

	Grade 8 (*n*=24)			
	M	*SD*	*Min*	*Max*
May 2012 EIa	220.79	38.06	166	287
September 2012 EIa	206.63	26.90	173	272
December 2012 EIa	211.58	25.38	164	272

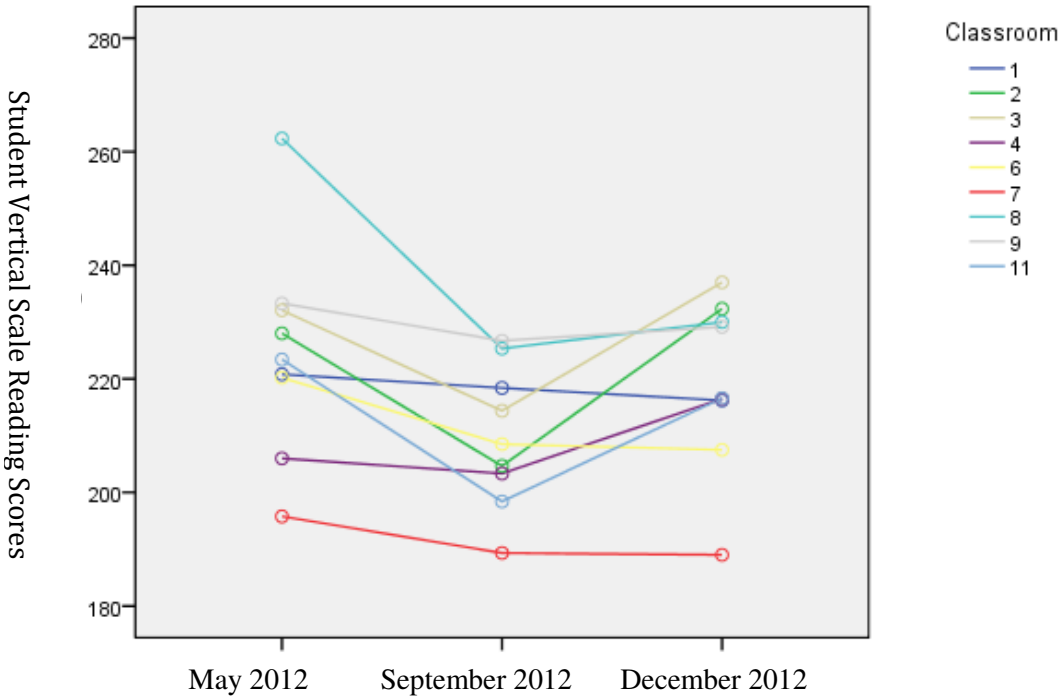
[a]Expressed as vertical scale scores.

Figure 3.1 Reading Performance Across Measured Time Points



A statistically significant relationship between student reading achievement and teachers' fidelity of implementation of the ELA curriculum overall was not established in the analyses of these data (question 1 of this study); however, further exploration of these data was conducted to examine if particular components of teachers' fidelity of implementation of the ELA curriculum had a statistically significant relationship with student reading achievement (question 2). To answer this question, the above-mentioned component elements of total fidelity were explored. Since adherence did not have any variability across participants (see Table 3.2) and dosage was uniquely related to the non-instructional variable (dosage = 100 – non- instruction) that left two main elements of fidelity of implementation of the ELA curriculum to be explored in relation to the EIa scores: quality of implementation and participant responsiveness (student engagement).

Figure 3.2 Reading Performance of Each Classroom across Timepoints



Quality of implementation (which contained the teacher non-instructional variable related to dosage) was explored first. In this category of fidelity of implementation, there were four main independent teacher behavior variables of interest. The main independent teacher behavior variables, namely, teaching function, feedback, non-instructional time (related to dosage), and quality instruction, were tabulated for each observation. In looking at these raw data, there was a lot of variability across observations. However, not all of the variability was of interest in this study. Observations varied both within and across teachers in many ways, including the length of time of each observation and the

Table 3.2

Means, Standard Deviations, and Ranges for Teacher Fidelity Variables[a]

Variable	*M*	*SD*	*Minimum*	*Maximum*
Total Fidelity	78.73	2.16	74.57	82.70
Adherence	100.00	0.00	100.00	100.00
Dosage [b]	68.06	15.64	28.41	95.60
Teaching Function (Overall)[c]	73.24	23.33	31.52	157.83
Advance Organizer	9.62	5.99	0.00	26.67
Review	12.81	16.21	0.00	76.39
Demonstrate	2.53	6.31	0.00	28.41
Model	1.33	6.20	0.00	38.71
Guided Practice	6.79	12.56	0.00	43.04
Peer Practice	7.80	19.21	0.00	68.13
Independent Practice	27.64	25.17	0.00	94.05
Formal Assessment	4.73	16.58	0.00	77.03
Non-Instruction (Overall)[c]	91.15	23.00	32.95	143.96
Transition	9.46	7.56	0.00	28.57
Anecdote	36.08	17.35	4.44	78.65
Monitoring	32.46	17.02	0.00	76.92
Off-Task	13.16	12.71	0.00	45.65
Feedback (Overall)[c]	67.21	24.90	31.87	119.32
Acknowledge Response	16.31	12.70	0.00	44.57
Elaborative Feedback	32.42	18.07	5.56	92.86
Questioning	18.47	14.89	0.00	57.95
Quality Indicators (Overall)[c]	119.03	55.07	3.57	196.59
Content Knowledge Instruction	4.80	13.33	0.00	56.52
Vocabulary Instruction	13.88	26.04	0.00	91.36

Means, Standard Deviations, and Ranges for Teacher Fidelity Variables

Variable	*M*	*SD*	*Minimum*	*Maximum*
Strategy Instruction	4.05	16.41	0.00	90.91
Discussion Reading Content	7.06	12.24	0.00	41.43
Motivation	4.68	5.09	0.00	19.44
Writing	25.55	34.81	0.00	97.78
Technology	59.02	33.98	0.00	100.00

[a]*Note*. *M*, Minimum, and Maximum expressed as percent of total observations.
[b]*Note*. Dosage represents an absence of non-instructional behavior; as such, it is related to the Non-Instructional variable (Dosage = 100 – NI).
[c]*Note*. *M* and Maximum may total greater than 100% since more than one component part of the overall element could occur simultaneously (e.g., Writing and Technology could co-occur in the Quality Indicators category).

type/topic of lesson taught. Variability due to the length of observation was not of interest to this study; I wanted to examine which teacher behaviors were associated with reading outcomes. Therefore, teachers were categorized into groups- those that were high on each variable and those who were low on each variable. To accomplish this, a median split was performed for each variable, as this created a natural break between teachers who engaged in high amounts of the activity captured by the variable, and those who engaged in lesser amounts of the activity captured by the variable. Examining the variables categorically to see if those teachers who engaged in the most amounts of the measured behaviors compared to those teachers who engaged in less amounts of the measured behaviors was a more effective way to investigate the impact of these behaviors on student reading scores without the nuisance interference of variability due to length of observation and other factors not of interest to this study. In creating the median spilt these data were recoded using a simple coding (0,1) system. All sums that fell below the median score for each

Table 3.3

Means and Standard Deviations for Teacher Fidelity Variables: Grade 6

	Grade 6					
	Teacher 1		Teacher 2		Teacher 3	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*
Total Fidelity	78.93	2.39	80.44	4.72	74.57	2.04
Teaching Function	66.47	7.81	62.61	25.24	77.40	12.35
Feedback	46.04	17.69	43.28	23.95	55.03	15.19
Non-Instruction	67.43	10.27	80.41	13.11	55.62	13.28
Quality	75.04	8.38	63.19	32.38	73.99	20.70
Student Engagement	85.76	4.93	85.00	7.86	73.85	10.88
	Teacher 4		Teacher 5		Teacher 6	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*
Total Fidelity	77.96	4.08	78.02	10.49	77.95	2.62
Teaching Function	66.21	20.24	71.21	11.46	90.99	3.24
Feedback	51.62	16.60	43.29	22.87	52.64	13.73
Non-Instruction	60.89	2.33	72.98	10.78	57.52	26.29
Quality	83.41	7.25	64.75	52.99	63.98	48.11
Student Engagement	83.86	4.10	79.33	21.40	85.08	10.16

Note. *M* expressed as percent of total observations.

Note. Due to the lack of variability in the variable, adherence is not reported. Furthermore, since dosage is related to the Non-Instructional variable (Dosage = 100 – NI), it is not reportedly separately from Non-Instruction.

variable were recoded into a score of 0, and all sums that were higher than the median score were recoded into a score of 1. This created a HI_LO variable for each of the initial independent teacher variables of interest. Thus, four new variables were created; namely, TEACHING FUNCTION HI_LO, FEEDBACK HI_LO, NON-INSTRUCTION HI_LO,

Table 3.4

Means and Standard Deviations for Teacher Fidelity Variables: Grade 7

	Grade 7			
	Teacher 1		Teacher 2	
Variable	*M*	*SD*	*M*	*SD*
Total Fidelity	82.34	6.78	76.90	1.01
Teaching Function	66.46	22.23	70.13	24.98
Feedback	45.08	18.78	40.05	18.17
Non-Instruction	77.85	6.46	67.31	18.56
Quality	83.46	24.40	95.04	3.89
Student Engagement	86.53	11.00	71.87	11.51

Note. *M* expressed as percent of total observations.

Note. Due to the lack of variability in the variable, adherence is not reported. Furthermore, since dosage is related to the Non-Instructional variable (Dosage = 100 – NI), it is not reportedly separately from Non-Instruction.

and QUALITY HI_LO. The transformed HI_LO independent teacher variables captured high levels of the variable compared to low levels of the variable (e.g., high numbers of occurrence of quality instruction vs. low numbers of occurrence of quality instruction, high numbers of occurrence of teaching functions vs. low numbers of occurrence of teaching functions). Since the observations occurred just after the final reading assessment was administered (final reading assessment occurred in December 2012; classroom observations occurred in January 2013), collected reading scores were compared with current observed teacher behaviors. By computing these transformed variables, one can make these direct comparisons between the collected reading scores and the independent variables recorded in the direct observations of teaching behaviors, for

Table 3.5

Means and Standard Deviations for Teacher Fidelity Variables: Grade 8

	Grade 8					
	Teacher 1		Teacher 2		Teacher 3	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*
Total Fidelity	74.91	2.94	82.70	4.25	78.89	3.80
Teaching Function	56.03	32.95	83.11	5.71	54.65	14.11
Feedback	47.37	14.51	42.64	9.09	47.58	9.01
Non-Instruction	61.73	20.71	83.20	12.48	73.11	7.60
Quality	70.26	23.24	77.46	20.88	62.50	22.25
Student Engagement	80.04	13.37	79.89	4.28	87.54	13.70

Means and Standard Deviations for Teacher Fidelity Variables: Grade 8

	Grade 8					
	Teacher 4		Teacher 5		Teacher 6	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*
Total Fidelity	76.97	3.78	78.93	1.15	71.57	7.27
Teaching Function	69.65	28.03	73.81	21.94	60.24	19.24
Feedback	50.07	3.30	51.22	36.18	51.60	9.98
Non-Instruction	71.63	1.32	62.34	23.75	60.86	26.07
Quality	43.94	33.74	94.50	7.69	7.69	8.21
Student Engagement	45.12	24.33	93.40	2.60	86.24	4.40

Note. *M* expressed as percent of total observations.

Note. Due to the lack of variability in the variable, adherence is not reported. Furthermore, since dosage is related to the Non-Instructional variable (Dosage = 100 – NI), it is not reportedly separately from Non-Instruction.

one can reasonably assume that the teaching observed in the January 2013 observations do not differ substantially from the teaching that occurred prior to January 2013. With these transformed variables, exploratory analyses could now be performed to determine which of the variables were most related to the student reading scores.

To compare the transformed independent teacher variables to the reading outcome measure, repeated measures ANOVAs were constructed to examine if any of the independent teacher variables were statistically significantly related to the student reading scores. As such, the TEACHING FUNCTION HI_LO, FEEDBACK HI_LO, NON-INSTRUCTION HI_LO, and QUALITY HI_LO variables were entered as a between subjects factor in a separate repeated measures ANOVAs. The NON-INSTRUCTION HI_LO teacher variable was found to have a statistically significant relationship to the EIa measure outcome measure ($F=12.197$, $p=.001$, $df=1, 54$). Analyses of the TEACHING FUNCTION HI_LO, FEEDBACK HI_LO, and QUALITY HI_LO variables failed to yield statistically significant outcomes, indicating that none of these variables could be identified as being related to students' reading performance. The plot of the relationship between the NON-INSTRUCTION HI_LO and EIa variables indicates a negative relationship between HI_LO non-instructional time and the EIa measure (see Figure 3.5). That is, reading scores increased as the time teachers spent engaged in non-instructional behavior decreased (SUMNIHI_LO=0.00). Conversely, as time teachers spent engaged in non-instructional behavior increased (SUMNIHI_LO=1.00), students' reading scores decreased. While this relationship is interesting, it is important to note that even though the analysis indicated that student reading scores were related to the non-instructional time

Figure 3.5 Reading performance at the three EIa Timepoints vs. HI_LO Non-Instructional Time

variable, differences existed between the students of teachers in the HI_LO groups at the first time point (May 2012). A t-test was performed to investigate the reading scores at the initial May 2012 time point to see if there were differences in the students' reading performance prior to students' assignment to their current classroom. The t-test indicated that there were statistically significant differences among the students in their reading performance prior to their assignment into their current classrooms with their current teachers (t=66.009, p<.001, df=93).

With the analysis of the quality of implementation variable's relationship to the reading EIa scores complete, next the relationship between the final variable of fidelity of implementation (namely, student engagement) and the student reading performance was examined. As described in the previous Methods chapter, in this study, three randomly selected students from among those in the sample in each classroom were observed using

a 30-second momentary time sampling (MTS) procedure on each observation. Students were coded to identify each interval during the observation in which they were clearly demonstrating off-task behavior. Thus, the student engagement variable actually provides information regarding the frequency with which each student was not engaged in instruction; it captured intervals of **off-task** behavior, as one can more reliably observe when someone is off-task rather than on-task. Thus, the student engagement variable actually captures student ***dis***engagement in teacher instruction. Since the observation periods had variable length in the total time of the observation, each student's average non-engaged time was calculated across the three separate observation sessions, yielding a variable labeled as Percent Off Student Engagement (OFFSE) for each observation.

As noted above in the quality of implementation variables, in looking at these raw data of the student engagement variable, there was a lot of variability across observations. However, not all of the variability was of interest in this study. Therefore, as in the above analyses with the quality of implementation variables, a median split was performed creating a natural break between students who spent larger amounts of time engaged in off-task behavior, and those who engaged in lesser amounts of off-task behavior. In creating the median spilt these data were recoded using a simple coding (0,1) system. All sums that fell below the median score for student disengagement were recoded into a score of 0, and all sums that were higher than the median score were recoded into a score of 1. This created a new variable named OFFSEHI_LO. With this transformed variable, the exploratory analyses to determine the relationship between student disengagement and the student reading scores could now be performed.

To compare the transformed student disengagement variable to the reading outcome measure, repeated measures ANOVAs were constructed to examine if it was statistically significantly related to the student reading scores. As such, the OFFSEHI_LO variable was entered as a between subjects factor in a repeated measures ANOVA. This analysis failed to yield statistically significant outcomes (F=1.474, p=.230, df=1, 59), indicating that student disengagement was not statistically significantly related to students' reading performance.

Research has demonstrated that student engagement can be strongly influenced by teachers' instructional behaviors (e.g., Cameron & Pierce, 1994; Deci, 1971; Wang & Holcombe, 2010), so student engagement was also compared to the independent teacher variables (teaching function, feedback, non-instructional behavior, and quality instruction) to examine if a relationship existed between these variables in the current study. Initially, these data were examined to establish if there was an identifiable and meaningful correlation between the variables. The correlation analyses indicated that there is a moderate correlation between student engagement and the quality instruction variable (r=-0.319, p=.001, moderate correlation; Cohen, 1988). This indicated that when no instances of quality instruction were noted on the part of the teacher, students were more disengaged. Conversely, when teachers spent more time engaged in quality instruction, student engagement increased. There was also a moderate correlation observed between student engagement and the teaching function variable (r=-0.306, p=.002). This indicated that when teachers did not engage in any of the teaching functions, students were more disengaged. Conversely, when teachers did engage in at least one teaching function (advance organizer, modeling, etc.), student engagement increased. None of the other two

main independent teacher variables (feedback and non-instructional behavior) had a significant correlation to student engagement. Specifically, the correlation between student engagement and feedback was r=.026 (p=.791) and the correlation between student engagement and teacher non-instructional time was r=.123 (p=.212).

While the correlations noted above were informative, the correlational analyses ignored the nested nature of the data (repeated measures nested within students who are nested within teachers). Therefore, these data were examined using hierarchical linear modeling (HLM v6.08 (Raudenbush & Bryk, 2002)) to investigate what proportion of the variability in student engagement scores could be explained by the teacher non-instructional behaviors.

Based on the observed correlations noted above, as well was the earlier results of the comparisons between the independent teacher variables and student reading performance, a subset of the original independent teacher variables were brought into HLM to be explored in comparison to the student engagement variable. First, a null model was explored to determine if there was, in fact, variability among the student engagement scores that could be modeled. Once the variability within the student engagement scores was verified with the null model, these data were explored to see if there was variability across time. Then the independent teacher variables were entered separately as time-varying covariates into the model. Finally, the independent teacher variables were transformed into aggregate categorical teacher characteristics and entered into the model at the teacher level to see if they could explain a greater proportion of variance than the null model.

Since the statistically significant correlations found in the previous analyses were between student engagement and teaching function and student engagement and quality instruction, these two variables were fully explored in HLM to examine which of these variables, or perhaps, which elements of these variables, could explain the variability in the student engagement scores. Because the teacher non-instructional variable produced statistically significant results when compared to student reading performance, it was important to look at this variable again in HLM where the nesting in these data could be accounted for in the analysis, despite the fact that the teacher non-instruction variable was not significantly correlated to student engagement in the previous analysis.

As mentioned in the previous chapter, the teacher variables all had subcomponent parts. Teaching function had eight subcomponent elements: advance organizer, review, demonstrate, model, guided practice, peer practice, independent practice, and formal assessment. Quality instruction had seven subcomponent elements: content knowledge, vocabulary, strategy instruction, discussion of reading content, motivation, technology, and writing. The teacher non-instructional variable had four component parts: transition, anecdote, monitoring, and off-task. As such, both the summary variables (teaching function, quality instruction, and teacher non-instruction) and their respective component parts were brought into HLM to be compared to the student engagement variable. Additionally, time was entered at level 1 to see if there was a linear relationship between time and student off-task behavior (e.g., did off task behavior increase across time). Time was captured by the Observation variable, which was simple coded so that observation at time 1 (observation with both observers present) was assigned a code of 0, observation at time two was assigned a code of 1 (single observer present), and observation at time 3

(single observer present) was assigned a code of 2. Descriptive statistics for variables used in these exploratory HLM analyses can be found in Table 3.6.

First, a one-way ANOVA with random effects (a fully unconditional null model) was analyzed to ensure there was enough detectable variability within the student engagement data to justify further exploring the relationships between student engagement and the teacher variables- both as time varying covariates and as aggregate teacher level variables. The general model is represented by the following equations:

Level-1: $Y_{tij} = \pi_{0ij} + e_{tij}$

Level-2: $\pi_{0ij} = \beta_{00j} + r_{0ij}$

Level-3: $\beta_{00j} = \gamma_{000} + u_{00j}$

Mixed Model: $Y_{tij} = \gamma_{000} + r_{0ij} + u_{00j} + e_{tij}$

With regards to this research study, Y_{tij} is the predicted student engagement score for time ***t*** for student ***i*** in teacher's classroom ***j***. The term π_{0ij} is mean student engagement for student ***i*** in teacher classroom ***j***. The final level-1 variable, e_{tij}, is the error term for level-1 (e.g., the difference between the student observed score at time *t* and the predicted score at time *t*). At level-2, β_{00j} is the mean student engagement for teacher classroom ***j*** and r_{0ij} is the error term (e.g., the difference between the student observed score and predicted score). Finally, at level-3, γ_{000} is the overall mean of student engagement across classrooms and u_{00j} is the level-3 error term (e.g., the difference between the teacher observed score and predicted score). This model assumes that at level-1, the error term is normally distributed with a mean of 0 and constant level-1 variance, $\boldsymbol{\sigma}2$ (Raudenbush &

Bryk, 2002). At level-2 and level-3, it is assumed that the error terms are multivariate normal with homogeneous variance. This model is fully unconditional in that no predictors are specified at level-1, level-2, or level-3. There is no test for statistical significance for variability at level 1 and is simply reported descriptively ($\sigma^2 = XXX$); the average student engagement score was positive and statistically significant (level 1; $\gamma_{000,}$= 5.8889, p<0.001, df= 13; see Table 3.7). There was statistically significant variability across teachers (level 3: $\tau_{\beta 00}$ = 2.12302, p=0.021, df= 13). However, this analysis failed to demonstrate statistically significant variability in student engagement scores between students within teacher classrooms (level 2). In other words, students in the same classroom demonstrated similar student engagement scores. Collectively, these results indicated that there was variability to be explained at level 3 (i.e., between teachers) and justified an exploratory analysis of the teacher non-instructional variable and its component parts to see how much of the variance any of these variables might be able to explain in these data.

The null model was then run again with the time predictor, the specific observation (observation 1, observation 2, or observation 3), to examine if the student off-task behavior was related to the specific observation (e.g., was there a linear trend between student off task behavior and time) The general model for the null equation with the time predictor is as follows:

Level-1: $Y_{tij} = \pi_{0ij} + \pi_{1ij}X_{tij} + e_{tij}$

Level-2: $\pi_{0ij} = \beta_{00j} + r_{0ij}$

$\pi_{1ij} = \beta_{10j} + r_{1ij}$

Table 3.6

Means, Standard Deviations, and Ranges for Variables Used in HLM Comparison of Teacher Non-Instructional Time and Student Engagement

	Level-1 Variables (*n*=126)			
Variable	*M*	*SD*	*Minimum*	*Maximum*
Student Engagement SUM	5.89	5.09	0.00	25.00
Teaching Function (TF) SUM	61.76	19.19	29.00	131.00
Advance Organizer SUM	8.10	5.00	0.00	24.00
Review SUM	10.69	12.97	0.00	55.00
Demonstrate SUM	2.19	5.46	0.00	25.00
Modeling SUM	1.19	5.68	0.00	36.00
Peer Practice SUM	6.26	15.35	0.00	62.00
Independent Practice SUM	23.71	21.75	0.00	79.00
Guided Practice SUM	5.64	10.11	0.00	34.00
Formal Assessment SUM	3.98	13.62	0.00	62.00
Quality Instruction (QI) SUM	94.12	47.81	3.00	176.00
Content Knowledge SUM	4.26	11.86	0.00	52.00
Vocabulary SUM	11.57	21.59	0.00	76.00
Strategy Instruction SUM	3.10	11.97	0.00	60.00
Discussion Reading Content SUM	5.93	10.26	0.00	37.00
Motivation SUM	3.86	4.14	0.00	16.00
Technology SUM	50.43	29.50	0.00	89.00
Writing SUM	21.71	29.71	0.00	88.00
Teacher Non-Instruction (NI) SUM	77.60	21.85	29.00	131.00
Transition SUM	8.12	6.62	0.00	26.00
Anecdote SUM	30.64	15.32	4.00	70.00
Monitoring SUM	27.48	14.53	0.00	70.00
Off-Task SUM	11.36	11.36	0.00	42.00

Means, Standard Deviations, and Ranges for Variables Used in HLM Comparison of Teacher Non-Instructional Time and Student Engagement

	Level-2 Variables (*n*=42)			
Variable	*M*	*SD*	*Minimum*	*Maximum*
Observation[a]	74.5	19.09	61.00	93.00

	Level-3 Variables (*n*=14)			
Variable	*M*	*SD*	*Minimum*	*Maximum*
Teaching Function AVERAGE	12.81	11.90	0.00	34.93
Review AVERAGE	27.64	18.44	4.12	71.11
Independent Practice AVERAGE	4.05	9.09	0.00	30.30
Teacher Non-Instruction (NI) AVERAGE	68.07	8.55	56.00	83.00
Transition AVERAGE	9.46	4.52	1.92	15.12

[a]*Note.* The Observation variable is expressed in number of 30-second intervals.

Level-3: $\beta_{00j} = \gamma_{000} + u_{00j}$

$\beta_{10j} = \gamma_{100} + u_{10j}$

Mixed Model: $Y_{tij} = \gamma_{000} + \gamma_{100} X_{tij} + r_{0ij} + r_{1ij} X_{tij} + u_{00j} + u_{10j} X_{tij} + e_{tij}$

In this model, the variables are as described above in the null model; the only exception is the addition of the time predictor variable, X_{tij}. Specifically, X_{tij} is the observation t in student i for teacher j. The addition of the time predictor to the null model equation did not yield a statistically significant result ($\gamma_{100,}$= -0.17, p=0.801, df= 13); thus, the student engagement scores were not linearly related to the specific observation (e.g., time).

Next, the data were explored to determine if the teacher variables and any of their component parts accounted for the variance in the student engagement scores at level-1 (that is, if the independent teacher variables, and/or any of their constituent parts, could

Table 3.7

HLM Descriptive Statistics for Level 1 Teacher Variables of Interest Models

	Model 1: Null			
Fixed Effects	Coefficient		*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})				
Intercept (γ_{000})	5.89	***	(0.582)	13
Random Effects	Variance		χ^2	*df*
Level 1				
Temporal variation, e_{tij}	23.60			
Level 2 (students within classrooms)				
Individual initial status, r_{0ij}	0.019		20.4719	28
Level 3 (between classrooms)				
Initial status, u_{00j}	2.123	**	25.3086	13

Note. ** indicates a statistically significant finding at the p=.05 level; *** indicates a statistically significant finding at the p<.001 level.

account for the variability among or were related to student engagement scores). The general model for these equations is as follows:

Level-1: $Y_{tij} = \pi_{0ij} + \pi_{1ij} W_{tij} + e_{tij}$

Level-2: $\pi_{0ij} = \beta_{00j} + r_{0ij}$

$\pi_{1ij} = \beta_{10j} + r_{1ij}$

Level-3: $\beta_{00j} = \gamma_{000} + u_{00j}$

$\beta_{10j} = \gamma_{100} + u_{10j}$

Mixed Model: $Y_{tij} = \gamma_{000} + \gamma_{100} W_{tij} + r_{0ij} + r_{1ij} W_{tij} + u_{00j} + u_{10j} W_{tij} + e_{tij}$

In these models, an independent teacher variables, W_{tij}, replaced the time predictor variable in the previous model. This time varying covariate W_{tij} was centered around the grand mean, and, on different iterations of the model, represented the over-arching teacher variables (NI, TF, and QI) as well as the constituent elements of these variables; such as, review, anecdote, vocabulary, etc. For each teacher variable entered into the equation (e.g., summary teaching function, modeling, etc.), the actual recorded score for each observation was entered as a level 1 predictor, similar to the time (observation) variable in the previous model. By doing so, one could examine whether student engagement varied as a function of teacher instructional behavior at the specific time at which the student behavior was observed. As such, twenty-two separate models were run, one model for each of the aforementioned variables, to investigate the relationship between student engagement and these teacher variables. Only one model, the one with Independent Practice entered at level 1 as a time-varying covariate, produced a statistically significant result (see Table 3.8 for results of fixed effects of the time-varying covariates at level 1 in all models run; note only models that met a priori criteria are reported). A statistically significant effect for the intercept in this model indicated that after controlling for independent practice, student engagement was positive; that is, as teachers engaged in more independent practice, the student engagement score was expected to increase. Thus, considering how the student engagement variable captures off-task behavior, students were more likely to be off-task when teachers employed more independent practice. The statistically significant main effect for the slope of the independent practice variable indicated that as teachers engaged in more independent

Table 3.8

HLM Descriptive Statistics for Level 1 Teacher Variables of Interest Models

	Model 3: Teaching Function (TF) SUM			
Fixed Effects	Coefficient		*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})				
Intercept (γ_{000})	5.8372	***	(0.598)	13
Model for teachers' teaching function behavior (β_{10j})				
Intercept (γ_{100})	0.0502	*	(0.025)	13
Random Effects	Variance		χ^2	*df*
Level 1				
Temporal variation, e_{tij}	22.2226			
Level 2 (students within classrooms)				
Individual initial status, r_{0ij}	0.0671		29.1973	28
Individual first-order school growth rate, r_{1ij}	0.0006		24.0693	28
Level 3 (between classrooms)				
Initial status, u_{00j}	2.4403	***	41.3440	13
First-order school growth rate, u_{10j}	0.0003	**	24.4825	13

Note. * p=.065; meets a priori criteria

practice that the student engagement score was expected to increase, e.g., students were more likely to be off task when teachers employed more independent practice.

Knowing that there was a small sample size at level-3 (at the teacher level; n=11), and that this could impact the ability to detect statistically significant results at level 1, and wanting to include all variables of interest when examining the variance at level 3, an *a priori* criteria was set such that any level 1 time-varying covariates that were less than or equal to the significance level of 0.1 ($p \leq .1$) would be included in the analyses at level 3.

HLM Descriptive Statistics for Level 1 Teacher Variables of Interest Models

	Model 4: Review SUM			
Fixed Effects	Coefficient		*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})				
Intercept (γ_{000})	5.8501	***	(0.555)	13
Model for teachers' review behavior (β_{10j})				
Intercept (γ_{100})	-0.0682	*	(0.038)	13
Random Effects	Variance		χ^2	*df*
Level 1				
Temporal variation, e_{tij}	22.7243			
Level 2 (students within classrooms)				
Individual initial status, r_{0ij}	0.0385		23.9116	28
Individual first-order school growth rate, r_{1ij}	0.0001		20.8815	28
Level 3 (between classrooms)				
Initial status, u_{00j}	1.5608		16.9313	13
First-order school growth rate, u_{10j}	0.0032		15.0252	13

Note. * p=.099; meets a priori criteria

Thus, five level 1 time-varying covariates (namely, teacher non-instruction SUM, teaching function SUM, review SUM, independent practice SUM, and transition SUM) were examined at level 3. Due to the statistically significant effect of Independent Practice at level 1, and the additional four teacher variables that *approached* statistical significance at level 1, one might predict that there was some variance that could still be accounted for within these data. Therefore, additional exploratory analyses were run to investigate whether or not any of the variance at level-3, the teacher level, could be explained. Due to

HLM Descriptive Statistics for Level 1 Teacher Variables of Interest Models

	Model 5: Teacher Independent Practice Behavior			
Fixed Effects	Coefficient		*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})				
Intercept (γ_{000})	5.8651	***	(0.455)	13
Model for teachers' independent practice behavior (β_{10j})				
Intercept (γ_{100})	0.1003	**	(0.029)	13
Random Effects	Variance		χ^2	*df*
Level 1				
Temporal variation, e_{tij}	15.602			
Level 2 (students within classrooms)				
Individual initial status, r_{0ij}	0.6224		19.3589	28
Individual first-order school growth rate, r_{1ij}	0.0084		36.0456	28
Level 3 (between classrooms)				
Initial status, u_{00j}	0.2577		14.2973	13
First-order school growth rate, u_{10j}	0.0588		17.1834	13

the suspected influence of the small sample size on the ability to detect effects at the teacher level noted above, the five variables (teaching function (TF) SUM, review SUM, Independent practice SUM, teacher non-instruction (NI) SUM, and transition SUM) were transformed for these analyses as was done earlier in the EIa ANOVA analyses. Since there was a notable range in the length of observation period between and among teachers, the five teacher variables were turned into an *average* rather than a sum variable. That is, to account for the variability in the sum variables, the sum of the number of intervals in which a teacher was engaged in the behavior of interest (e.g., review) was averaged over

HLM Descriptive Statistics for Level 1 Teacher Variables of Interest Models

	Model 6: Teacher Non-Instruction SUM			
Fixed Effects	Coefficient		*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})				
Intercept (γ_{000})	5.7194	***	(0.613)	13
Model for teachers' transition behavior (β_{10j})				
Intercept (γ_{100})	0.0506	*	(0.030)	13
Random Effects	Variance		χ^2	*df*
Level 1				
Temporal variation, e_{tij}	18.662			
Level 2 (students within classrooms)				
Individual initial status, r_{0ij}	0.1548		20.8542	28
Individual first-order school growth rate, r_{1ij}	0.0027		25.5668	28
Level 3 (between classrooms)				
Initial status, u_{00j}	2.9734	**	31.9546	13
First-order school growth rate, u_{10j}	0.0055	**	25.0074	13

Note. * p=.116; meets a priori criteria

total observation periods (across total observed intervals) to eliminate the variability due to length of observation, which was not of interest to this study. Once the variable was averaged for each observation for each teacher, a summary average was computed to represent the teacher's average TF, NI, review, independent practice, and transition score across all three observations. These new variables were named teacher non-instruction AVERAGE (AVENI), teaching function AVERAGE (AVETF), review AVERAGE (AVEReview), independent practice AVERGAE (AVEIndPrac), and transition AVERAGE (AVETRANS). Once these average variables were computed, a median split was performed, as this created a natural break between teachers who engaged in high

HLM Descriptive Statistics for Level 1 Teacher Variables of Interest Models

	Model 7: Teacher Transition Behavior			
Fixed Effects	Coefficient		*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})				
Intercept (γ_{000})	5.8775	***	(0.532)	13
Model for teachers' transition behavior (β_{10j})				
Intercept (γ_{100})	0.1213	*	(0.071)	13
Random Effects	Variance		χ^2	*df*
Level 1				
Temporal variation, e_{tij}	23.4208			
Level 2 (students within classrooms)				
Individual initial status, r_{0ij}	0.0148		26.3014	28
Individual first-order school growth rate, r_{1ij}	0.0002		17.9766	28
Level 3 (between classrooms)				
Initial status, u_{00j}	1.3451	**	34.5991	13
First-order school growth rate, u_{10j}	0.0012	**	24.6344	13

Note. * p=.112; meets a priori criteria

amounts of these behaviors, and those who engaged in lesser amounts of these behaviors. In creating the median spilt these data were recoded using a simple coding (0,1) system. All averages that fell below the median score for each variable were recoded into a score of 0, and all averages that were higher than the median score were recoded into a score of 1. This created five additional variables, named AVENIHI_LO which captured the median split of the average NI variable, AVETFHI_LO which captured the median split of the average TF variable, AVEReviewHI_LO which captured the median split of the average review variable, AVEIndPracHI_LO which captured the median split of the average independent practice variable, and AVETRANSHI_LO, which captured the median split

of the average transition variable. This allowed for examination of the effect for teachers who overall engaged in comparatively higher amounts of the variable of interest to teachers who overall engaged in comparatively lower amounts of the variable of interest. These new variables were then explored using a three-level HLM model to examine the fixed effects and the proportion of variance in the student engagement scores that might be able to be explained by these variables. The transformed aggregate HI_LO variables were entered as predictors at level 3 in the models. The general model for these equations is as follows:

Level-1: $Y_{tij} = \pi_{0ij} + e_{tij}$

Level-2: $\pi_{0ij} = \beta_{00j} + r_{0ij}$

Level-3: $\beta_{00j} = \gamma_{000} + \gamma_{001}D_j + u_{00j}$

Mixed Model: $Y_{tij} = \gamma_{000} + \gamma_{001}D_j + r_{0ij} + u_{00j} + e_{tij}$

In these models, an independent teacher non-instructional variable, D_j, was added as a predictor to the model at level-3. D_j represented the median split HI_LO variables, on the five separate iterations of the above model. Using the formula below[1], the results of these analyses indicated that the AVENIHI_LO variable accounts for almost 50% of the variance in teachers at level-3 (see Table 3.9 for results). Fixed effects were not statistically significant for any of the above models.

---

[1] The variance between teachers in the intercept: $\frac{\tau_{\beta 00 unrestricted} - \tau_{\beta 00 full}}{\tau_{\beta 00 unrestricted}} = \frac{2.12-1.13}{2.12} = .466$

Table 3.9

HLM Descriptive Statistics for Aggregate Teacher Variables Models

	Model 8: Average Overall Non-Instruction HI_LO		
Fixed Effects	Coefficient	*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})			
Intercept (γ_{000})	6.7500 ***	(0.6858)	12
Model for teacher behavior effect (β_{10j})			
Intercept (γ_{100})	-2.0093	(1.0476)	12
Random Effects	Variance	χ^2	*df*
Level 1			
Temporal variation, e_{tij}	23.5988		
Level 2 (students within classrooms)			
Individual initial status, r_{0ij}	0.0186	20.4719	28
Level 3 (between classrooms)			
Initial status, u_{00j}	1.1343	20.0423	12

HLM Descriptive Statistics for Aggregate Teacher Variables Models

	Model 9: Average Transition Behavior HI_LO			
Fixed Effects	Coefficient		*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})				
Intercept (γ_{000})	5.2698	***	(0.790)	12
Model for teacher behavior effect (β_{10j})				
Intercept (γ_{100})	1.2381		(1.117)	12
Random Effects	Variance		χ^2	*df*
Level 1				
Temporal variation, e_{tij}	23.5988			
Level 2 (students within classrooms)				
Individual initial status, r_{0ij}	0.0186		20.4719	28
Level 3 (between classrooms)				
Initial status, u_{00j}	1.7398	***	23.2674	12

HLM Descriptive Statistics for Aggregate Teacher Variables Models

	Model 10: Average Overall Teaching Function HI_LO			
Fixed Effects	Coefficient		*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})				
Intercept (γ_{000})	6.3969	***	(0.8012)	12
Model for teacher behavior effect (β_{10j})				
Intercept (γ_{100})	-1.0159		(1.0159)	12
Random Effects	Variance		χ^2	*df*
Level 1				
Temporal variation, e_{tij}	23.5988			
Level 2 (students within classrooms)				
Individual initial status, r_{0ij}	0.0186		20.4719	28
Level 3 (between classrooms)				
Initial status, u_{00j}	1.8650	**	23.9344	12

HLM Descriptive Statistics for Aggregate Teacher Variables Models

	Model 11: Average Review Behavior HI_LO		
Fixed Effects	Coefficient	*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})			
Intercept (γ_{000})	5.1588 ***	(0.776)	12
Model for teacher behavior effect (β_{10j})			
Intercept (γ_{100})	1.4603	(1.330)	12
Random Effects	Variance	χ^2	*df*
Level 1			
Temporal variation, e_{tij}	23.5988		
Level 2 (students within classrooms)			
Individual initial status, r_{0ij}	0.0186	20.4719	28
Level 3 (between classrooms)			
Initial status, u_{00j}	1.5899 **	22.4689	12

HLM Descriptive Statistics for Aggregate Teacher Variables Models

	Model 12: Average Independent Practice HI_LO		
Fixed Effects	Coefficient	*(SE)*	*df*
Model for Intercept of student engagement (β_{00j})			
Intercept (γ_{000})	5.7444 ***	(0.6855)	12
Model for teacher behavior effect (β_{10j})			
Intercept (γ_{100})	0.5056	(1.2825)	12
Random Effects	Variance	χ^2	*df*
Level 1			
Temporal variation, e_{tij}	23.5988		
Level 2 (students within classrooms)			
Individual initial status, r_{0ij}	0.0186	20.4719	28
Level 3 (between classrooms)			
Initial status, u_{00j}	2.0709 **	25.0308	12

Chapter 4

DISCUSSION AND IMPLICATIONS

As many scholars have noted, the educational system in the United States is entrenched in a literacy crisis (e.g., Boardman et al., 2008; Haynes, 2005; Moje et al., 2008; Torgesen et al., 2007). While this literacy crisis effects all ages and grade levels of students, a population of students that warrants particular concern is adolescent readers. Lee, Grigg, and Donahue (2007) found that approximately two-thirds of both eighth- and twelfth-graders read below proficiency and lack the reading skills needed to succeed in school.

With the national movement towards a Response to Intervention (RTI) context of identifying and intervening with struggling learners, the Tier I general education classroom is the first line of defense in preventing and intervening with literacy difficulties. Proponents of RTI believe that comprehensive and coordinated instruction that is implemented with fidelity can improve outcomes for all students. However, fidelity has not been clearly defined in the literature, and has historically received less attention in the K-12 education literature than in other fields (such as the health field) (Ruiz-Primo, 2005; Summerfelt, 2003). Fortunately, the literature regarding quality instructional practices is relatively well defined, and general consensus exists as to what quality instruction should entail in adolescent literacy classrooms (Colvin, Flannery, Sugai, & Monegan, 2009).

In the current study, an observational tool was constructed and used to record the fidelity of implementation middle school teachers of English/Language Arts (ELA) classes employed over several observations of their teaching. Observational data were

compared to student reading performance to examine relationships between fidelity and student reading performance. Results indicated relationships between dosage and reading performance as well as between student engagement and independent practice.

Discussion of Results

This section describes the analysis of comparisons between student reading performance and measures of teachers' fidelity of implementation to the ELA curriculum in adolescent literacy classrooms. First, teachers' overall fidelity of implementation is compared to student reading performance, and then subcomponent parts of fidelity of implementation are compared to student reading performance to explore relationships between these variables. Then, variables of interest were brought into a hierarchical linear modeling program (HLM v6.08; Raudenbush & Bryk, 2002) to examine the proportion of variance accounted for by teacher variables related to fidelity of implementation.

Overall fidelity of implementation. Teachers' overall fidelity of implementation to the ELA curriculum (calculated as equal parts adherence, dosage, quality of implementation and student engagement) was not found to have a statistically significant effect on student reading achievement. Several factors need to be considered, however, when interpreting this finding. First, given the small sample size at level-3 (teacher level; n=11), it was difficult to determine the existence of statistically significant relationships among these data. The small sample size reduced the statistical power of the data. Second, the literature does not currently provide guidance with regards to how one should conceptualize *overall* fidelity of implementation to core curriculum when you compute fidelity of implementation from component parts. That is, it is unclear whether all elements of fidelity of implementation contribute equally to overall fidelity of

implementation; as such, giving each component equal weight in calculations of overall fidelity of implementation may not be the best approach to determining an individual's overall fidelity of implementation to a core curriculum. Some aspects, such as dosage, may be more important than other aspects (such as adherence, which lacked variability in this study). Going forward, it will be important to determine which elements of fidelity of implementation are crucial and which may be supplemental to more effectively determine how to weight components when calculating overall fidelity of implementation (Noell, 2008).

Adherence. There was not any variability in teachers' adherence to the ELA curriculum observed in this study. As such, this component of fidelity was not found to have a statistically significant effect on student reading achievement. While this did not result in a statistically significant finding, it is notable that all teachers demonstrated complete adherence to the ELA curriculum. Since adherence accounted for 25% of teachers' overall fidelity of implementation to the ELA curriculum, it is possible that the lack of variability on this element of fidelity of implementation contributed to the lack of detectable statistically significant differences among teachers in their overall fidelity of implementation to the ELA curriculum. This effectively worked to reduce the potential variability among teachers by one-quarter, or 25%. It is possible that if there had been variability on this measure of fidelity of implementation, statistically significant findings may have been detectable among teachers in their overall fidelity of implementation to the curriculum. From an educational standpoint, however, this lack of variability due to 100% adherence to the curriculum, is quite encouraging. When school districts develop their curriculum, they intend for it to be delivered in its entirety and with quality instructional

behaviors. Complete adherence among all teachers in the study over the observational period provides promising evidence that the district is providing sufficient training and support to enable teachers to deliver the curriculum as written.

Dosage. Students in classrooms with teachers who spent more time engaged in delivering the ELA curriculum (as opposed to engaged in off-task or non-instructional behaviors such as managing classroom behavior) demonstrated statistically significant higher reading performance than students in classrooms with teachers who spent less time engaged in delivering the ELA curriculum. Differences in dosage among teachers accounted for almost 50% of the variance among teachers in this study. Thus, even though *what* the teachers taught (adherence) was found to lack variability among teachers, and all teachers were found to completely adhere to the ELA curriculum, there was significant variability in the *amount of time* teachers spent implementing the curriculum in their classrooms. Again, since there was complete adherence to the ELA curriculum, this variability in dosage is largely related to time spent transitioning between activities, time spent engaged in off-task activities, or time spent engaged in non-instructional activities, such as managing student behavior. Given this finding, attention should be given to providing supports to teachers to ensure that they are able to make the most effective use of their instructional time. Providing strategies to better manage instructional time, including the management of student behavior, may be helpful in allowing teachers to increase the dosage of curriculum provided to students. Since the data in this study demonstrate a statistically significant relationship between increased dosage and increased student reading performance, we must make strides to ensure that students are in classrooms with teachers who maximize the dosage of the curriculum. Given the noted

literacy issues in this country, we should feel compelled to ensure that our adolescents are enrolled in classrooms with teachers who implement the ELA curriculum with high dosage to improve students' reading performance.

Quality of Implementation. No elements of quality of implementation were found to have a statistically significant effect on student reading achievement. When examining the data related to quality of implementation, however, several of the component parts of quality of implementation (such as content knowledge instruction, strategy instruction, and motivation) occurred at such low rates (about 6% of the total observations), it is likely that these behaviors were not observed to occur frequently enough to yield detectable results. Elements that occurred with greater frequency (such as technology and writing) tended to occur in all-or-nothing manners; that is, when a teacher used technology, they tended to use it across observations and for the duration of an observation, thereby reducing the variability *within* teachers on these component parts of this construct. Therefore, it is possible that these difficulties in observing the component parts of quality of implementation contributed to the non-statistically significant findings in this aspect of fidelity of implementation to the ELA curriculum.

Student Engagement. Student engagement was not found to have a statistically significant effect on student reading achievement. To ensure that an appropriate number of observations were conducted to get a reliable and valid estimate of student engagement, only three students were observed per classroom. Even though these students were selected at random from those students successfully recruited into the study in each classroom, it is possible that those students selected at random were not representative of all students in the class in each classroom. Collecting data on the engagement of a larger

number of students in each classroom may add additional insight into the relationship that may exist between student engagement and reading performance.

Student engagement was found to have a statistically significant relationship with another independent variable examined in this study- independent practice. Time spent in independent practice (a teaching function included in quality of implementation) influenced student engagement; as teachers engaged in more independent practice, students were more likely to be off-task. When teachers engage in the independent practice teaching function, they typically can only provide feedback or direction instruction to one student at a single time. Therefore, since teacher attention is generally limited to a specific student during this teaching function, it is possible that this creates an opportunity for more students to become off-task or disengaged, as the teacher has a more limited ability to monitor the entire class for engagement.

Implications

The major findings of this study were that students in classrooms where teachers demonstrated higher dosages of the ELA curriculum had statistically significantly higher reading performance than students in classrooms with teachers who demonstrated low levels of dosage of the ELA curriculum. This finding may provide some insight into which elements of fidelity of implementation (adherence, dosage, quality, student engagement, program differentiation) are more critical than other elements of this construct. Even if further studies do not extend this finding that dosage has a differential effect on overall fidelity of implementation that other elements of this construct, practical implications for this finding still exist. Teachers need to be provided with strategies (which may include management of student behavior, decreasing time spent transitioning

between classroom activities, etc.) to assist them in improving the amount of time they spend engaged in instruction (dosage). The evidence obtained in this study indicates that increased dosage increases student reading performance, so given the literacy crisis our country currently faces, addressing the dosage one devotes to the ELA curriculum for adolescent readers should be an educational priority.

Moreover, this study also demonstrated that student engagement was statistically significantly related to time spent in independent practice. Students in classrooms where they engaged in higher amounts of independent practice were less engaged than students in classrooms where teachers spent time in other teaching functions, such as reviewing, guided practice, etc. Even though this study failed to demonstrate a statistically significant relationship between student engagement and reading performance, the literature generally supports the idea that students who are more engaged in instruction experience greater academic success (e.g., Fredricks, Blumfeld, & Paris, 2004). Independent practice is an important element of the process by which students gain self-regulated mastery over skills and knowledge that they are learning (e.g., Archer & Hughes, 2011; Biancarosa & Snow, 2006; Coyne et al., 2009; Guthrie et al., 2007; Mastropieri, Scruggs, & Graetz, 2003; Rosenshine, 1995; Rosenshine, 1997; Rosenshine & Stevens, 1986; Swanson, 1999; Swanson & Deshler, 2003). As such, it is important to support teachers in creating effective instructional routines within their classrooms so that when students are provided with opportunities to practice literacy skills independently, student engagement does not decrease. If students are not engaged in the instructional activity during independent practice, then they are not positioned to benefit from the opportunity to develop and enhance their self-regulated mastery of the learning material.

Time spent by teachers early in the school year establishing classroom procedures and routines has been demonstrated to result in increased engagement in academic tasks by students throughout the rest of the school year (Cameron, Connor, & Morrison, 2005; Evertson & Emmer, 1982). Thus, assisting teachers in developing such routines and reinforcing the notion that time spent establishing these procedures has significant beneficial effects on student achievement, is an important consideration for teacher preparation and initial mentoring.

Limitations

This section notes the limitations of this study that should be considered when interpreting the results and implications of the study. Limitations related to the observational tool developed, the statistical power of the study, and the representativeness of the sample are described below.

Observational Tool. The tool developed for this study to collect the observations of teacher fidelity of implementation to the ELA curriculum effectively captured the component elements of fidelity of implementation relevant to this study (namely, adherence, dosage, quality of implementation, and participant responsiveness). However, while the literature has been clear on what constitutes quality instruction in ELA classrooms (Colvin, Flannery, Sugai, & Monegan, 2009), some aspects of this instruction do not lend themselves well to detection through time-sampling procedures. Furthermore, those that can be detected through time-sampling procedures do not lend themselves well to being observed simultaneously with other behaviors. For example, Mastropieri, Scruggs, and Graetz's 2003 review of research on reading comprehension instruction for adolescents found that when teachers provide students with clear directions and strategies,

students increase their understanding of what they read. This type of explicit instruction is clearly an element of quality instruction that should be considered when measuring this element of fidelity of implementation to ELA curriculum for adolescent readers. Defining "clear" directions is something of a subjective exercise, and as such, does not lend itself well to being observed through time-sampling procedures with numerous other variables that are being simultaneously observed. As such, it is possible that this element of fidelity of implementation was not fully defined and explored in this study, which may limit the interpretations and conclusions that can be drawn from this analysis.

Statistical Power. This study did not attain the minimal sample size needed in order to examine the relationships among the study variables with the appropriate level of statistical power. A priori recruitment targets for student participants was set at 210 students to demonstrate significant findings with the power of alpha set at 0.05 and a medium effect size of 0.4. The actual recruited sample of students was well below this target (n=126). While the actual sample size was able to detect statistically significant results for the dosage variable, it is possible that the small sample size limited the detection of statistically significant findings for the other variables of interest in this study. Furthermore, the small sample size at level-3 (teacher level; n=11) reduced the statistical power of the data at this level of analysis as well. As such, it was difficult to determine the existence of statistically significant relationships among these data, particularly when modeling the different equations that attempted to account for the nesting of the data in this study.

Representativeness of Sample. The sample of students utilized in this study had statistically significant differences in their reading performance prior to their assignment

into their current classrooms with their current teachers (t=66.009, p<0.001, df=93). This a priori difference in students may have had an influence on the noted relationship between teacher dosage of the ELA curriculum and student reading performance. That is, it is possible that teachers of students who started with lower reading scores spent more time managing student behavior or attending to other non-instructional tasks, thereby maintaining the differences in the reading scores. Since these differences among the students existed prior to their assignment to their ELA teacher, it is possible that the different abilities among the students at baseline also contributed to the variability in teacher dosage of the ELA curriculum.

Suggestions for Future Research

The results of this study indicated that elements of fidelity of implementation to ELA curriculum are related in a statistically significant manner to student reading performance, and that time spent in independent practice decreases student engagement in literacy instruction. The limitations associated with this study suggest interesting avenues for future research to build upon these findings and increase the literature related to fidelity of implementation to core ELA curriculums for adolescent readers. Suggestions for modifications to the current study to assist in future replications of this research are described below.

Sample size and sampling concerns. This study did not attain the minimal sample size needed in order to examine the relationships among the study variables with the appropriate level of statistical power. Thus, future research should focus upon replicating this study with an appropriate sample of both adolescent readers (approximately 210 students) and teachers of adolescent ELA curriculum (approximately 15 different

teachers). In order to improve future studies' ability to examine student engagement in the literacy instruction, attempts should be undertaken to secure permission to videotape the classroom so that all students can be recorded and coded for engagement at a later time. Not only would this enable a greater number of students to be observed, but it would also allow observers to more accurately record teaching behaviors and more efficiently resolve disagreements in observational coding. Further, future replications should attend to a priori differences in student reading performance and attempt to control for these differences in both study procedures and analyses.

Dependent measure of students' reading performance. This study utilized school district-developed Error Identification Assessments (EIa), which are diagnostic reading assessments that similar to Cognitive Diagnostic Assessments (CDAs). The literature supports the use of EIa to reliably serve a predictive function for future reading performance, provide a lens through which teachers can identify students who are at-risk for failing to meet established district reading performance benchmarks, and provide targeted intervention to assist students in improving their reading performance. The error analysis element of the EIa allows teachers to more accurately select intervention tools to target areas of weakness in students' development of reading skills, as well as inform their general teaching practices. However, the EIa used for this study, as noted earlier, are district-developed instruments, and as such, may lack appropriate psychometric properties to reliably measure student reading performance. Future research should focus upon validating this measure of reading performance, or, alternatively, utilizing measures of student reading performance with proven reliability and validity.

Instrumentation. The observational tool developed for this study was not able to effectively capture all of the known elements of quality instruction (e.g., explicit instruction). Future research should attend to the development of an observational tool that easily and objectively measures all relevant aspects of quality of implementation. Without all facets of quality of implementation appropriately represented in the observational tool utilized to collect the fidelity of implementation data, it is possible that important aspects of this construct will be overlooked or under-represented in the data.

Determining overall fidelity of implementation. Currently, the fidelity of implementation literature related to core curriculums does not provide guidance with regards to how one should conceptualize *overall* fidelity of implementation to core curriculum when you compute fidelity of implementation from component parts. That is, it is unclear whether all elements of fidelity of implementation contribute equally to overall fidelity of implementation. Future investigations into fidelity of implementation to core ELA curriculums should focus upon determining which elements of fidelity of implementation are crucial and which may be supplemental (Noell, 2008). This information will provide critical guidance for determining how to weight components when calculating overall fidelity of implementation to core curriculums.

Summary

Aspects of fidelity of implementation to Tier I ELA curriculum are meaningfully related to student reading achievement. The dosage, or time teachers spent delivering the curriculum, was related to statistically significantly higher reading achievement scores among students in classrooms where teachers displayed higher levels of dosage of the

ELA curriculum. Furthermore, student engagement in instruction, an element of fidelity of implementation, was found to be influenced by time spent in independent practice. While these results are promising, there were some methodological limitations to this study that may have influenced the obtained results. Specifically, the small sample size at level-3 (teacher level; n=11) reduced the statistical power of the data. As such, it was difficult to determine the existence of statistically significant relationships among the data. Additionally, there were statistically significant differences among the students in their reading performance prior to their assignment into their current classrooms with their current teachers (t=66.009, p=.000, df=93). Thus, it is possible that teachers of students who started with lower reading scores spent more time managing student behavior or attending to other non-instructional tasks, thereby maintaining the differences in the reading scores. Since these differences among the students existed prior to their assignment to their ELA teacher, it is possible that the different abilities among the students at baseline also contributed to the variability in teacher dosage of the ELA curriculum.

Future research conducted with larger sample sizes that control for differences among student reading achievement at baseline will help determine if additional aspects of fidelity of implementation to Tier I ELA curriculum exist that may have been suppressed in this study. The relative contributions of each separate element of fidelity of implementation (adherence, dosage, quality of implementation, and student engagement) should be investigated to clarify if all of these components contribute equally to one's fidelity of implementation to Tier I ELA curricula.

References

ACT. (2005). Reading between the lines: What the ACT reveals about college readiness in reading. Iowa City, IA: Author.

Adelman, C. (2006). The toolbox revisited: Paths to degree completion from high school through college. Washington, DC: U.S. Department of Education.

Albee, G. W., & Gullotta, T. [Eds.]. (1997). *Primary prevention works*. Thousand Oaks, CA: Sage.

American Educational Research Association; American Psychological Association; National Council on Measurement in Education (1999). *Standards of educational and psychological testing*. Washington, DC: American Education Research Association.

Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal, 40,* 685-730.

Archer, A. L., & Hughes, C. A. (2011). *Explicit instruction: Effective and efficient teaching*. New York: Guilford Press.

Ariail, M. & Albright, L. (2006). A survey of teachers‘ read-aloud practices in middle schools. *Reading Research and Instruction, 45*(2), 69-89.

Armento, B. J. (1976). *Teacher behaviors related to student achievement on a social science concept test* (Doctoral Dissertation). Retrieved from Dissertation Abstracts International. (Accession No. 1979-02158-001).

Baumann, J. F., Edwards, E. C., Boland, E. M., Olejnik, S, & Kame'enui, E. J. (2003) Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meanings. *American Education Research Journal, 40,* 447-494.

Baumann, J. F., Edwards, E. C., Font, G., Tereshinski, C. A., Kame'enui, E. J., & Olejnik, S. (2002). Teaching morphemic and contextual analysis to fifth-grade students. *Reading Research Quarterly, 37*, 150-176.

Bear, D. R., Invernizzi, M., Templeton, S., & Johnston, F. (2007). *Words Their Way: Word study for phonics, vocabulary, and spelling instruction* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Biancarosa, G., & Snow, C. E. (2006). *Reading next- A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York (2nd ed.).* Washington, DC: Alliance for Excellent Education.

Boardman, A. G., Roberts, G., Vaughn, S., Wexler, J., Murray, C. S., & Kosanovich, M. (2008). *Effective instruction for adolescent struggling readers: A practice brief.* Center on Instruction.

Buehl, D. (2008). *Classroom Strategies for Interactive Learning*. International Reading Association.

Burns, M. S., Griffin, P., & Snow, C. E. (Eds.). (1999). *Starting out right: A guide to promoting children's reading success*. National Academies Press.

Cameron, C. E., Connor, C. M., & Morrison, F. J. (2005). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology, 43,* 61-86.

Cameron, J., & Pierce, W.D. (1994). Reinforcement, reward, and intrinsic

motivation: A meta-analysis. *Review of Educational Research, 64(3*), 363–424.

Cantrell, S. C., Almasi, J. F., Rintamaa, M., Carter, J. C., Pennington, J., & Buckman, D. M. (2014). The Impact of Supplemental Instruction on Low-Achieving Adolescents' Reading Engagement. *Journal Of Educational Research*, *107*(1), 36-58.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007) A conceptual framework for implementation fidelity. *Implementation Science, 2(40).* Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2213686/

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (second ed.). Lawrence Erlbaum Associates.

Collins, J. J. (2007). *The Collins writing program: Improving student performance through writing and thinking skills across the curriculum*. West Newbury, MA: Collins Education Associates.

Colvin, G., Flannery, K. B., Sugai, G., & Monegan, J. (2008). Using observational data to provide performance feedback to teachers: A high school case study. *Preventing School Failure, 53*, 95-104.

Connecticut State Department of Education. (2008). *Using scientific research-based interventions: Improving education for all students*. Middletown, CT: Connecticut State Department of Education. Retrieved from http://www.sde.ct.gov/sde/lib/sde/pdf/pressroom/SRBI_full.pdf.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied Behavior Analysis* (2nd ed.). pper Saddle River, New Jersey: Pearson.

Cornett, J. (2010). *What's evidence got to do with it? an observational study of research-based instructional behavior in high school classes.* (Order No. 1476574, University of Kansas). *ProQuest Dissertations and Theses,* , 60. Retrieved from http://ezproxy.lib.uconn.edu/login?url=http://search.proquest.com/docview/496944391?accountid=14518. (496944391).

Coyne, M. D., Zipoli, JR., R. P., Chard, D. J., Faggella-Luby, M., Ruby, M., Santoro, L. E., & Baker, S. (2009). Direct instruction of comprehension: Instructional examples from intervention research on listening and reading comprehension. *Reading & Writing Quarterly, 25*, 221-245. Doi: 10.1080/10573560802683697

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology*, *83*(2), 264.

Curtis, M. (2004). Adolescent who struggle with word identification: Research and practice. In T. L. Jetton, & J. A. Dole (Eds.), *Adolescent literacy research and practice* (pp. 119-134). New York: Guildford.

Cushing, L. S., & Kennedy, C. H. (1997). Academic effects of providing peer support in general education classrooms on students without disabilities. *Journal of Applied Behavior Analysis*, *30*(1), 139-151.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, *18*, 23–45.

Deci, E.L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18(1)*, 105–115. doi:10.1037/h0030644

Domitrovich, C. E., Bradshaw, C. P., Poduska, J., Hoagwood, K. E., Buckley, J., Olin, S. S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion: Training and Practice, Research and Policy, 1,* 6-28.

Dobson, L. D., & Cook, T. J. (1980). Avoiding Type III error in program evaluation: Results from a field experiment. *Evaluation and Program Planning*, *3*, 269–276.

Doyle, W. (1986). Classroom organization and management. *Handbook of research on teaching*, *3*, 392-431.

Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. *What research has to say about reading instruction*, *3*, 205-242.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American journal of community psychology*, *41*(3-4), 327-350.

Erlich, O., & Shavelson, R. J. (1978). The Search for Correlations between Measures of Teacher Behavior and Student Achievement: Measurement Problem, Conceptualization Problem, or Both?. *Journal of Educational Measurement*, *15*(2), 77-89.

Ehren, B. J. (2005). Looking for evidence-based practice in reading comprehension instruction. *Topics in Language Disorders, 25,* 310-321.

Englert, C. S., & Thomas, C. C. (1987). Sensitivity to text structure in reading and writing: A comparison between learning disabled and non-learning disabled students. *Learning Disability Quarterly*, *10*(2), 93-105.

Evertson, C. M., & Emmer, E. T. (1982). Effective management at the beginning of the school year in junior high classes. *Journal of Educational Psychology, 74,* 485-498.

Faggella-Luby, M. N., Ware, S. M., & Capozzoli, A. (2009). Adolescent literacy-Reviewing adolescent literacy reports: Key components and critical questions. *Journal of Literacy Research, 41,* 453-475.

Faggella-Luby, Schumaker, and Deshler, 2009

Feuer, M.J., Holland, P.W., Green, B.F., Bertenthal, M.W., & Hemphill, F.C. (Eds.) (1999). *Uncommon measures: Equivalence and linkage among educational tests.* Washington, DC: National Research Council.

Fitzgerald, J., and Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist, 35*, 39–50.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature.* Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication No. 231). Retrieved August 18, 2010, from http://cfs.fmhi.usf.edu/_docs/publications/NIRN_Monograph_Full.pdf

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.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice, 18,* 157-171.

Gajria, M., Jitendra, A. K., Sood, S., & Sacks, G. (2007). Improving comprehension of expository text in students with ld a research synthesis. *Journal of learning disabilities*, *40*(3), 210-225.

Gambrell, L. B. (2011). Seven rules of engagement: What's most important to know about motivation to read. *The Reading Teacher, 65,* 172-178.

Gardenier, N. C., MacDonald, R., & Green, G. (2004). Comparison of direct observational methods for measuring stereotypic behavior in children with autism spectrum disorder. *Research in Developmental Disabilities, 25*, 99-118.

Garner, R. (1985). Text summarization deficiencies among older students: Awareness or production ability? *American Educational Research Journal, 22,* 549-560.

Graham, S., & Hebert, M. A. (2010). *Writing to read: Evidence for how writing can improve reading.* A Carnegie Corporation Time to Act Report. Washington, DC: Alliance for Excellent Education.

Graves, M. F., Juel, C., & Graves, B. B. (2004). Reading instruction for English language learners. *Teaching reading in the 21st century*, 496-541.

Gump, P. V. (1967, February). *Setting variables and the prediction of teacher behavior*. Paper presented to the American Educational Research Association, New York.

Guthrie, J. T., Anderson, E., Alao, S., & Rinehart, J. (1999). Influences of concept-oriented reading instruction on strategy use and conceptual learning from text. *Elementary School Journal, 99,* 343-366.

Guthrie, J. T., & McCann, A. D. (1997) Characteristics of classrooms that promote motivations and strategies for learning. In J. T. Guthrie & A Wigfield (Eds.), *Reading engagement: Motivating readers through integrated instruction* (pp. 128-148). Newark, DE: International Reading Association.

Guthrie, J. T., McRae, A., & Klausa, S. L. (2007). Contributions of Concept-Oriented Reading Instruction to knowledge about interventions for motivations in reading. *Educational Psychologist, 42,* 237-250.

Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology, 92,* 331-341.

Harris, R. E., Marchand-Martella, N., & Martella, R. C. (2000). Effects of a peer-delivered corrective reading program. *Journal of Behavioral Education, 10*, 21–36.

Hattie, J. (1999, August). Influences on student learning. *Inaugural lecture given University of Auckland*.

Hattie, J. (2005). What is the nature of evidence that makes a difference to learning?. *Australian Council for Education Research-Using data to support learning*, 7.

Haynes, M. (2005). *Reading at risk: How states can respond to the crisis in adolescent literacy*. Alexandria, VA: National Association of State Boards of Education, 2005.

Henderlong, J., & Lepper, M. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin, 128,* 774-795.

Henry, M. K. (2003). *Unlocking literacy: Effective decoding and spelling instruction*. Baltimore, MD: Brookes.

Hock, M. F., & Deshler, D. D. (2003). Adolescent literacy: Ensuring that no child is left behind. *Principal Leadership, 13(4)*, 55-61.

Ivey, G. (2003). The intermediate grades: The teacher makes it more explainable and other reasons to read-aloud in the intermediate grades. *The Reading Teacher*, *56*(8), 812-814.

Jacobs, V. A. (2008). Adolescent literacy: Putting the crisis in context. *Harvard Educational Review*, *78*(1), 7-39.

Jang, E. E. (2008). A framework for cognitive diagnostic assessment. In C. A. Chapelle, Y.-R. Chung, & J. Xu (Eds.), *Towards adaptive CALL: Natural language processing for diagnostic language assessment* (pp. 117-131). Ames, IA: Iowa State University.

Jenkins, J. R., Matlock, B., & Slocum, T. A. (1989). Two approaches to vocabulary instruction: The teaching of individual word meanings and practice in deriving word meaning from context. *Reading Research Quarterly, 14,* 379-396.

Jones, H. A., Clarke, A. T., & Power, T. J. (2008). Expanding the concept of intervention integrity: A multidimensional model of participant engagement. *Balance*, *23*, 4-5.

Kamil, M. L. (2003). Adolescents and literacy: Reading for the 21st century. Washington, DC: Alliance for Excellent Education.

Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A Practical Guide* (NCEE #2008-4027). Washington, DC: National Center for

Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc.

Kennedy, M. M. (1999). Approximations to indicators of student outcomes. *Educational Evaluation and Policy Analysis*, *21*(4), 345-363..

Kolen, M.J. & Brennan, R.L. (2004). *Test equating, scaling & linking, methods and practices (2nd edition).* New York: Springer.

Larson, A.F. & Kelleher, M. E. (2009, April). *Error identification assessments: Diagnostic instruments for instruction*. Paper presented at the American Educational Research Association's Annual Meeting, San Diego, CA.

Lee, J., Grigg, W., & Donahue, P. (2007). The nation's report card. *Reading*, 2007-496.

Maheady, L., Sacca, M. K., & Harper, G. F. (1987). Classwide student tutoring teams: The effects of peer-mediated instruction on the academic performance of secondary mainstreamed students. *The Journal of Special Education, 21(3),* 107-121.

Maheady, L., Sacca, M. K., & Harper, G. F. (1988). Classwide peer tutoring with mildly handicapped high school students. *Exceptional Children*.

Malone, L. D., & Mastropieri, M. A. (1991). Reading comprehension instruction: Summarization and self-monitoring training for students with learning disabilities. *Exceptional Children, 58,* 270-279.

Marzano, R. J., Association for Supervision and Curriculum Development, & Heinle/Cengage Learning. (2010). *Teaching basic and advanced vocabulary: A framework for direct instruction*. Heinle.

Mastropieri, M. A., Scruggs, T. E., & Graetz, J. E. (2003). Reading comprehension instruction for secondary students: Challenges for struggling students and teachers. *Learning Disability Quarterly, 26,* 103- 116.

Moats, L. (2001). When older students can't read. *Educational Leadership, 58(6),* 36-46.

Moje, E. B., Peek-Brown, D., Sutherland, L. M., Marx, R., Blumenfeld, P. and Krajcik, J. 2004. "Explaining explanations: Developing scientific literacy in middle-school project-based science reforms". In D. Strickland & D.E. Alvermann (Eds.), *Bridging the gap: Improving literacy learning for preadolescent and adolescent learners in grades 4–12* (pp.227–151). New York: Teachers College Press.

Mowbray, C., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, *24*, 315–340.

Munby, H., & Russell, T. (1990). Metaphor in the study of teachers' professional knowledge. *Theory Into Practice*, *29*, 116–121.

Murphy, G., & Goodall, E. (1980). Measurement error in direct observations: A comparison of common recording methods. *Behavior Research and Therapy, 18*, 147-150.

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology, 101,* 740-764.

National Assessment Governing Board. (2007). *Reading framework for the 2009 National Assessment of Educational Progress.* Washington, DC: American Institutes for Research.

National Center for Education Statistics. (2005). Mapping 2005 state proficiency standards onto the NAEP scales (NCES 2007-482). Washington, DC: U.S. Department of Education.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.

National Institute of Child Health and Human Development (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, DC: U. S. Government Printing Office.

No Child Left Behind (NCLB) Act of 2001, 20 U.S.C.A. § 6301 et seq. (West 2003).

Noell, G. H. (2008). Research examining the relationships among consultation process, treatment integrity, and outcomes. In W. P. Erchul & S. M. Sheridan (Eds.), *Handbook of research in school consultation: Empirical foundations for the field* (pp. 323-342). Mahwah,NJ: Lawrence Erlbaum Associates.

Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L., et al. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review, 34,* 87-106.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78,* 33-84.

Ozgungor, S., & Guthrie, J. T. (2004). Interactions Among Elaborative Interrogation, Knowledge, and Interest in the Process of Constructing Knowledge From Text. *Journal of Educational Psychology*, *96*(3), 437.

Powell, J., Martindale, B., Kulp, S., Martindale, A., & Bauman, R. (1977). Taking a closer look: Time sampling and measurement error. *Journal of Applied Behavior Analysis, 10*, 325-332.

Power, T. J., Eiraldi, R. B., Clarke, A. T., Mazzuca, L. B., & Krain, A. L. (2005). Improving mental health service utilization for children and adolescents. *School Psychology Quarterly*, *20*(2), 187.

Ransford, C. R., Greenberg, M. T., Domitrovich, C. E., Small, M., & Jacobson, L. (2009). The role of teachers' psychological experiences and perceptions of the curriculum supports on the implementation of a social and emotional learning curriculum. *School Psychology Review, 38,* 510-532.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, *75*, 211–246.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32,* 155-175.

Rissman, L. M., Miller, D. H., & Torgesen, J. K. (2009). *Adolescent literacy walk-through for principals: A guide for instructional leaders.* Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Rosenshine, B. V. (1980). Skill hierarchies in reading comprehension. *Theoretical issues in reading comprehension*, 535-554.

Rosenshine, B. (1995). Advances in research on instruction. *The Journal of Educational Research, 88,* 262-268.

Rosenshine, B. (1997). The case for explicit, teacher-led, cognitive strategy instruction. *MF Graves (Chair), What sort of comprehension strategy instruction should schools provide*.

Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd ed.) (pp. 376-391). New York: MacMillan Publishing Company.

Routman, R. (1991). *Invitations changing as teachers and learners K-12*. Portsmouth, NH: Heinemann Educational Books, Inc.

Ruiz-Primo, M. A. (2005, April). A multi-method and multi-source approach for studying fidelity of implementation. In S. Lynch (Chair) & C. L. O'Donnell, *"Fidelity of implementation" in implementation and scale-up research designs: Applications from four studies of innovative science curriculum materials and diverse populations*. Symposium conducted at the annual meeting of the American Educational Research Association, Montreal, Canada.

Sáenz, L. M., & Fuchs, L. S. (2002). Examining the reading difficulty of secondary students with learning disabilities expository versus narrative text. *Remedial and Special Education*, *23*(1), 31-41.

Salinger, T. (2011). *Addressing the "crisis" in adolescent literacy.* Paper prepared for the U.S. Department of Elementary and Secondary Education, Smaller Learning Communities Program, Herndon, VA.

Sanetti, L. M., & Kratochwill, T. R. (2009). *Introduction to the special series: Toward developing a science of treatment integrity: Implications for research, practice, and policy*. *School Psychology Review, 38,* 445-459.

Schunk, D. H., & Rice, J. M. (1992). Influence of reading-comprehension strategy information on children's achievement outcomes. *Learning Disability Quarterly, 15,* 51-64.

Sencibaugh, J. M. (2005, April). *A synthesis of content enhancement strategies for teaching students with learning problems at the secondary level.* Paper presented at the Council for Exceptional Children International Conference, Baltimore, MD.

Shavelson, R., & Dempsey-Atwood, N. (1976). Generalizability of measures of teaching behavior. *Review of Educational Research*, 553-611.

Shkedi, A. (1998). Can the curriculum guide both emancipate and educate teachers? *Curriculum Inquiry*, *28*, 209–231.

Snow, C. E., Porche, M. V., Tabors, P. O., & Harris, S. R. (2007). *Is literacy enough?* Baltimore, MD: Brookes.

Snowling, M. J., & Hulme, C. (Eds.). (2005). *The science of reading: A handbook.* Malden, MA: Blackwell Publishing.

Songer, N. B., & Gotwals, A. W. (2005, April). Fidelity of implementation in three sequential curricular units. In S. Lynch (Chair) & C. L. O'Donnell, *"Fidelity of implementation" in implementation and scale-up research designs: Applications*

from four studies of innovative science curriculum materials and diverse populations. Symposium conducted at the annual meeting of the American Educational Research Association, Montreal, Canada.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56,* 72-110.

Sterling-Turner, H. E., Watson, T. S., & Moore, J. W. (2002). The effects of direct training and treatment integrity on treatment outcomes in school consultation. *School Psychology Quarterly, 17,* 47-77.

Summerfelt, W. T. (2003). Program strength and fidelity to evaluation. *Applied Developmental Science*, *7*(2), 55–61.

Swanson, H. L. (1999). Instructional components that predict treatment outcomes for students with learning disabilities: Support for a combined strategy and direct instruction model. *Learning Disabilities Research & Practice*, *14*, 129-140.

Swanson, H. L., & Deshler, D. (2003). Instructing adolescents with learning disabilities: Converting a Meta-Analysis to Practice. *Journal of Learning Disabilities, 36,* 124-135.

Templeton, S. (2004). The vocabulary-spelling connection: Orthographic development and morphological knowledge at the intermediate grades and beyond. In J. Baumann & E. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp. 118-138). New York: Guilford.

Torgesen, J. K., Houston, D. D., Rissman, L. M., Decker, S. M., Roberts, G., Vaughn, S., ... & Lesaux, N. (2007). Academic literacy instruction for adolescents: A guidance

document from the Center on Instruction. *Portsmouth, NH: RMC Research Corporation, Center on Instruction*.

Towne, L., Wise, L. L., & Winters, T. M. (Eds.), for the Committee on Research in Education, National Research Council. (2005). *Advancing scientific research in education*. Washington, DC: National Academies Press.

U.S. Department of Education, Office of Communications and Outreach. (2010). *Overview of the U.S. Department of Education*, Washington, D.C.

U.S. Department of Education. (2006). *Education research grant request for application: CFDA Number 84.305*. Washington, DC: Institute for Education Sciences. Retrieved August 18, 2010, from http://ies.ed.gov/ncer/funding.

Walberg, H. J. (1986). Synthesis of research on teaching. *Handbook of research on teaching*, *3*, 214-229.

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61,* 620-630.

Wang, M., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal, 47(4)*, 633–662. doi:10.3102/0002831209361209

Wexler, J., Edmonds, M., & Vaughn, S. (2008). *Effective reading strategies for adolescent struggling readers.* Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth or their reading. *Journal of educational psychology*, *89*(3), 420.

Winograd, P. N. (1984). Strategic difficulties in summarizing texts. *Reading Research Quarterly*, 404-425.

Wright, C. J., & Nuthall, G. (1970). Relationships between teacher behaviors and pupil achievement in three experimental elementary science lessons. *American educational research journal*, 477-491.

Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Counseling & Clinical Psychology, 49,* 156-167.

Appendix B

Date:_____ **Start Time:_____** **End Time:_____**

Lesson Topic:_____

	Transition Time																							
Engagement	Student Not Engaged																							
	Off-Task																							
	Anecdote																							
	Monitoring/Rdg																							
Teaching Function	Review																							
	Present- Adv. Org.																							
	Present- Demonstrate																							
	Present- Modeling																							
	Peer Practice																							
	Guided Practice																							
	Independent Practice																							
	Acknowledge Student Response																							
	Elaborative Feedback																							
	Formal Assessment																							
Quality Indicators	Questioning																							
	Content Knowledge Instr.																							
	Vocabulary																							
	Strategy																							
	Discussion Rdg Content																							
	Motivation																							
	Technology																							
	Writing																							

Appendix C

Code Book for Classroom Observation Tool

Transition Time		
Description of Category	Transition time refers to instructional time that represents time moving from one activity, place, lesson, etc. to another.	
Relevant Sources	Doyle, 1986; Walberg, 1986	
Quality Indicator	**Operational Definition**	**Examples**
Transition Time	Transition time is class time that is not devoted to instruction, often because instruction has not yet begun (students have not yet settled into class) or because the teacher is moving the students from one activity to another. Transition time is not necessarily *unengaged* time, as the teacher may still be engaged, in that they are directing students in their transition.	Teacher directs student(s) to: • Take out learning materials (text book, paper, calculator, etc.) • Change seats (e.g., move into groups for discussion, collaboration, or assistance) • Put away learning materials (text book, paper, calculator, etc.) • Shift attention from one activity to another (e.g., stop reading and begin writing, stop writing and begin reading, stop discussing and begin reading, stop discussing and begin listening to lecture, etc.) Transition time may also occur: • As students enter the classroom and settle in for instruction

Engagement*		
Description of Category	Engagement refers to the actions, behaviors, and interactions that the teacher/student(s) partake in to demonstrate their involvement with instruction.	
Relevant Sources	Anderson & Walberg, 1994; Haynes & Jenkins, 1986	
Quality Indicator	**Operational Definition**	**Examples of Teacher** Behavior**
Student Not Engaged	Student engagement is defined as the actions, behaviors, and interactions the students are expected to engage in when participating in the delivery of the curriculum. However, since engagement is difficult to measure through observation, this study will determine whether students are *unengaged*, or *off-task*, at the given interval. If the student is *not* observed to be unengaged in instruction, then it is assumed that the student is engaged in instruction.	Unengaged students may: • Engage with non-instructional social talk with peers • Use their cell phone • Violate rules • Be out of the classroom (e.g., in the restroom) • Play a non-instructional game (either on paper or on the computer) • Involved in getting organized (e.g., getting out materials, putting materials away) • Throw something away • Do nothing (including not doing the teacher-assigned task)

Quality Indicator	Operational Definition	Examples of Teacher** Behavior
Off-Task (Non-interaction, non-instruction)	Teacher engagement is defined as the actions, behaviors, and interactions the teacher is expected to engage in when delivering the curriculum. However, since engagement is difficult to measure through observation, this study will determine whether the teacher is *unengaged*, or *off-task*, at the given interval. If the teacher is *not* observed to be unengaged in instruction, then it is assumed that the teacher is engaged in instruction.	Unengaged teachers may: • Talk with another adult (about something other than instruction (conversing with a co-teacher to troubleshoot instruction shall NOT be coded in this category)) • Talk on the phone • Write a student pass • Complete paperwork or computerized forms • Read professional reading materials • Access, write, or send emails • Engage in personal activities (e.g., reading a newspaper, filing nails, etc.)
Monitoring/Reading (Non-interaction instruction)	Check student work in progress (circulates room), reads aloud to students	• Check homework completion • Grade papers • Pass out papers • Take attendance
Anecdote (Interaction, non-instruction)	Engages students in anecdotes, etc. that can be tangentially related to topic	• Discussion with students not grounded in text; engaged in anecdote-telling • Managing student behavior

* Adopted from Cornett, 2010

***Note: In the case of Student Engagement, STUDENT Behavior is described.*

Teaching Function		
Description of Category	Teaching function refers to behaviors and procedures teachers engage in to deliver instruction to and assess learning of students.	
Relevant Sources	Archer & Hughes, 2011; Biancarosa & Snow, 2006; Coyne et al., 2009; Guthrie et al., 2007; Mastropieri, Scruggs, & Graetz, 2003; Rosenshine, 1995; Rosenshine, 1997; Rosenshine & Stevens, 1986; Swanson, 1999; Swanson & Deshler, 2003	
Quality Indicator	**Operational Definition**	**Examples**
Review	Any instruction in which the teacher reviews previous learning.	• Reviews homework or previous assignments • Reviews previously taught strategy • Reviews previously read/discussed text
Advance Organizer: Presentation	Previews what students are about to learn/do, gives a rationale for learning, sets goals, and/or gives directions (*prior to* engaging in instruction around this new content, strategy, text, etc.)	• Describes text students will read • Describes strategy students will learn • Describes activity students will engage in • Describes information students will learn • States learning goals of lesson

Demonstration: Presentation	Any instruction in which the teacher engages in delivering the intended lesson. This type of teaching function involves the delivery of *new* learning, whether the entire presentation is new, or the teacher is presenting a new aspect of something previously learned.	• Communicates new material
		  ○ Lecture
		  ○ Demonstration
		• Presents skills/tasks in small steps/pieces
		• Provides examples and non-examples

Quality Indicator	Operational Definition	Examples
Modeling: Presentation	Any activity in which the teacher shows and tells the students what to do. The teacher works through an example of what he/she wants the students to do on their own, by simultaneously demonstrating and telling students what to do to complete the activity successfully. The teacher may also demonstrate this instructional approach by "thinking aloud" as they show students how to use/apply a comprehension strategy. The teacher guides the students through the skill, strategy, example, etc., gradually decreasing scaffolding.	• Models or states how to perform a task When modeling how to use a graphic organizer (such as a *Know-Want to Know-Learned* chart), the teacher may use the think-aloud strategy: • Beginning with the K-Know column, the teacher may say to the students, "One thing that I know about the Holocaust is that it occurred during World War II. I am going to write this fact down in our K-Know column of our chart. Does anyone know anything else about the Holocaust that we can add to this column of the chart?"
Peer Practice	Any instructional arrangement in which the students are working collaboratively in small peer groups (of at least 2 students), as directed by the teacher.	*NOTE: In this learning arrangement, the teacher's instruction, when present, is aimed at student peer pairs. During this learning arrangement, the teacher may move from group-to-group to check progress and understanding but does not devote his/her instructional time exclusively to any one particular group of students.* The teacher may assign the students to: • Discuss reading content • Engage in peer writing conferences • Collaborate in research

Quality Indicator	Operational Definition	Examples
Guided Practice	Any instruction in which the teacher is involved with leading students through learning. The teacher actively assists students in practicing the new learning.	*NOTE: A distinguishing characteristic that separates Guided Practice from Teacher Modeling is that, in Guided Practice, the **students** are performing the work (not the teacher).* Supports students in engaging in learning activities (student participates in learning with scaffolded support from the teacher) by guiding students through a learning activity. The teacher may initially begin a response and have the students complete the answer; or, the teacher may provide a prompt, or "hint", such as: • When prompting students in the steps of finding the main idea, the teacher may say, "Remember, after naming the person/thing that the paragraph is about, the next step is to say what the person/thing did in all of the sentences." • The teacher provides students with a prompt, guiding them to select or use an appropriate cognitive strategy to assist in their comprehension of text.

Quality Indicator	Operational Definition	Examples
Independent Practice	Any instruction in which the students are directed by the teacher to work independently on an assignment (without assistance from any instructor or peer).	The teacher may assign the students to: • Read an assigned passage from the current reading selection • Read for pleasure • Read to gain information (research) • Work independently on a worksheet • Work independently on a quick, in-class assignment • Begin assigned homework • Work independently on a writing task that will go through the writing process of planning, drafting, revising, editing, publishing, etc. (such as a story, poem, short or extended research report, etc.) • Work independently on any written assignment that will take more than one day to complete
Acknowledge Student Responses	The teacher briefly recognizes student response.	• The teacher may respond verbally by saying, "Yes," "No," "Good job" or another short acknowledgment. • The teacher may respond nonverbally by nodding head, shaking head, giving a thumb's up, etc.

Quality Indicator	Operational Definition	Examples
Elaborative Feedback	The teacher provides feedback to student responses.	• If the student responds correctly, the teacher provides positive reinforcement. • If the student responds incorrectly, the teacher immediately corrects the error. *Note: Error correction should be carried out in a positive manner that builds the student's self-efficacy rather than creating a sense of failure.*
Formal Assessment	A formal assessment shall be considered any long (test; at least ½ of the class period) or short (quiz; several minutes to less than ½ of the class period) assessment that measures student knowledge for the purpose of assigning students a grade for their performance.	Examples of tests include: • End-of-unit summative assessment • Weekly vocabulary test • State/district-wide assessment (e.g., Connecticut Academic Performance Test (CAPT)) Examples of quizzes include: • Daily vocabulary quiz • Assessment of learning of previous day's instruction (e.g., retention of facts, strategies, procedures, etc.)

Quality Instruction

Description of Category	Quality instruction refers to the effective, research-supported instructional procedures teachers use to deliver curriculum to adolescent readers.
Relevant Sources	Applebee et al., 2003; Archer & Hughes, 2011; Baumann et al., 2003; Baumann et al., 2002; Bear, Invernizzi, Templeton, & Johnston, 2007; Biancarosa & Snow, 2004; Buehl, 2009; Coyne et al., 2009; Cunningham & Stanovich, 1991; Curtis, 2004; Duke & Pearson, 2002; Ehren, 2005; Englert & Thomas, 1987; Faggella-Luby, Schumaker, & Deshler, 2009; Fitzgerald & Shanahan, 2000; Gajria et al., 2007; Gambrell, 2011; Garner (1985); Graham & Hebert, 2010; Graves, Juel, & Graves, 2004; Guthrie & McCann, 1997; Guthrie et al., 1999; Guthrie et al., 2000; Guthrie et al., 2007; Henderlong & Lepper, 2002; Henry, 2003; Jenkins, Matlock, & Slocum, 1989; Kamil et al., 2008; Marzano (2010); Mastropieri, Scruggs, & Graetz, 2003; Moats, 2001; Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2007; NICHD, 2000; Ozgungor & Guthrie (2004); Reznitskaya et al., 2001; Rosenshine, 1995; Saenz & Fuchs, 2002; Schunk & Rice, 1992; Snow, Porche, Tabors, & Harris, 2007; Stahl & Fairbanks, 1986; Swanson, 1999; Swanson & Deshler, 2003; Templeton, 2004; Wexler, Edmonds, & Vaughn, 2008; Wigfield & Guthrie, 1997; Winograd (1984)

Quality Indicator	Operational Definition	Examples
Questioning	Questioning refers to the instructional practice of providing instructional cues in the form of questions to convey elements of content to be learned or directions for learning.	• Asks students frequent questions to check their understanding (e.g., checks to ensure students understand assigned task)

Quality Indicator	Operational Definition	Examples
Content Knowledge Instruction	Content knowledge instruction refers to instruction devoted to increasing students' knowledge of the content area. Teachers engage in content knowledge instruction by building upon and clarifying what students already know, and introducing new declarative, procedural, and conditional knowledge.	• Assess what students already know about the topic prior to instruction • Provide a short review of what the students have already learned about this topic, clarifying any student inaccuracies • Relate what the students will be learning to something they already know • Present facts (such as the names of characters, important dates, etc.), concepts, procedures (, such as the skills or steps involved in a process or strategy), or conditional information (such as when (and *when not*) to use a skill/strategy or other information)
Word Reading/ Vocabulary Instruction	Word reading instruction refers to teaching in which teachers provide students with instruction regarding how to read, understand, and use unknown words.	• Instructing students in recognizing word parts, such as root words, affixes, and/or syllabication • Instructing students in how to use available reference materials to derive the meaning of unfamiliar words. • The teacher introduces/teaches a word to increase students' general vocabulary knowledge or domain-specific knowledge

Quality Indicator	Operational Definition	Examples
Strategy Instruction	Comprehension strategy instruction refers to teaching in which teachers provide students instruction regarding how to use specific reading strategies to improve comprehension of text.	• The teacher instructs students in the use of a single reading comprehension strategy, such as: comprehension monitoring, summarizing/paraphrasing, question asking/generating, knowledge of text structures/features, graphic organizers, text structures, text features, inferencing, etc. • The teacher instructs students in the use of a packaged reading comprehension strategy, such as: Reciprocal Teaching, POSSE, Transactional Strategies Instruction, Collaborative Strategic Reading, etc.

Discussion of Reading Content	Discussion of reading content refers to teaching in which teachers engage students (either with the teacher or with others in the classroom) in discussion of the content of text.	• The teacher models, instructs, or engages students in:
		  ○ using others' questions and comments to build discussion
		  ○ expressing opinions or taking a position
		  ○ making connections across time and subjects
		  ○ questioning the author
		• The teacher asks students authentic questions, questions that are used to explore and develop knowledge rather than "test" student knowledge
		  ○ What do you think…?
		  ○ What else might you say…?
		  ○ How might this be viewed from _____'s perspective?
		  ○ What might happen if _____?
		  ○ In what other way might we show/illustrate _____?

Quality Indicator	Operational Definition	Examples
Motivation	Motivation refers to instructional behaviors exhibited by the teacher that foster students' motivation to learn and engage in instruction.	• Focus students on important and interesting learning goals • Provide a range of activity choices • Provide interesting texts at multiple reading levels • Provides connection between learning and relevance to students' lives

Technology	The teacher uses a technological device to enhance and/or delivery literacy instruction in the classroom.	There are a variety of technological tools a teacher may use, including:
		• Computer
		  ○ To explore websites
		  ○ To assist students in typing written products
		  ○ To assist students in creating multimedia presentations
		  ○ To deliver supplemental instruction (e.g., a literacy intervention)
		• Digital projectors
		  ○ To project PowerPoint slides that accompany instruction
		  ○ To display images
		  ○ To display exemplars of assignments
		• Smart board
		  ○ In lieu of a white board or chalkboard to present written information to students
		• Camera
		• Video recorder
		• Audio recorder
		• Audio player, including MP3 player

Quality Indicator	Operational Definition	Examples
Writing	Writing refers to teaching in which teachers engage students in activities that utilize their writing skills to improve their comprehension of text.	• The teacher provides opportunity for students to write about what they read (e.g., write a response to a text, write a summary of a text, answer questions in writing about a text) • The teacher engages in instruction devoted to increasing students' knowledge of the skills and processes needed to create text (e.g., spelling, sentence construction, text structures for writing)