Teacher Implementation of an Adolescent Reading Intervention

Margaret Troyer

Nonie K. Lesaux
James S. Kim
Catherine E. Snow

A Thesis Presented to the Faculty
of the Graduate School of Education of Harvard University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Education

2017
Acknowledgments

This work was supported by:

- the Institute of Education Sciences, U.S. Department of Education, through Grant R305F100026 to the Strategic Education Research Partnership as part of the Reading for Understanding Research Initiative. The opinions expressed are those of the author and do not represent views of the Institute or the U.S. Department of Education.

- the International Literacy Association under the Steven A. Stahl Research Grant

- Harvard University through a Dean’s Summer Fellowship

I would like to extend my personal gratitude to:

- the four teachers who allowed me to visit their classrooms and gave their precious time to answering my questions

- the members of my committee, Nonie Lesaux, Jimmy Kim, and Catherine Snow, as well as Lowry Hemphill, for their support throughout the process

- my writing group, Katie Leech and Bonnie Mackintosh, and the READS lab, for providing feedback on various drafts

- my colleague Mary Burkhauser for assistance with coding

- my parents, Mary O’Connell and Terry Troyer, for everything they have done to get me here, but most especially providing childcare so I could work

- my husband, James Haas, for supporting me throughout the process

- my daughter, Eleanor Troyer-Haas, for forcing me to work when it was work time and not work when it was playtime
Abstract

This dissertation examines teacher implementation of an adolescent literacy intervention with a coaching component, guided by questions about fidelity of implementation (FoI) and curriculum adaptation. In the first of two studies, I used data from observations of teachers (n=17) in nine schools during the 2013-14 school year to conduct a nuanced descriptive analysis of FoI. I also analyzed weekly logs completed by literacy coaches (n=3) to examine variation in quantity and intensity of coaching. I then compared variation in coaching with variation in FoI, and finally compared FoI to outcomes for students (n=287). FoI at observation 1 was found to predict coaching time, and FoI across both observations predicted student outcomes. This emphasizes the critical role of investigating implementation in order to better understand the results of intervention research.

In the second study, I used qualitative methodology to analyze adaptations made by four experienced teachers in one school that sustained implementation of this curriculum after the intervention trial had ended. Six focal adaptations were identified, three each from two teachers, and analyzed for productivity using criteria from Debarger and colleagues (Debarger, Choppin, Beauvineau, & Moorthy, 2013). Of the six, only two met criteria for productivity. This suggests that making productive adaptations is difficult, and that teachers should be supported to do so through educative curriculum materials and effective professional development.

In addition, an account of teaching practice methodology was used to define each teacher’s orientation toward the curriculum (Simon & Tzur, 1999), and then to determine whether this orientation demonstrated assimilation or accommodation to intervention
principles (Coburn, 2004). I found that the vast majority of time spent implementing the curriculum included adaptations, and that each teacher’s adaptations were different. Although one teacher demonstrated assimilation and accommodation to intervention principles, the other three primarily demonstrated assimilation. These findings suggest the importance of understanding teachers’ orientations toward curriculum in order to provide more tailored professional development which may help teachers accommodate to the most critical pedagogical features of a curriculum.
Chapter 1.

Introduction

Two-thirds of American eighth graders cannot proficiently read and comprehend text at grade level (National Center for Education, 2015). On average, African-American and Latino students enter high school with literacy skills three years behind those of Asian and White students, and students from low-income families enter high school with literacy skills five years behind those from high-income families (Reardon, Valentino, & Shores, 2012). Clearly, something must be done to improve adolescents’ literacy outcomes – especially in schools serving low-income students and students of color.

Many promising interventions have been designed to improve adolescents’ reading comprehension, but when implemented by classroom teachers, the effects of these interventions are often small (e.g., Cantrell, Almasi, Carter, Rintamaa, & Madden, 2010; Faggella-Luby & Wardwell, 2011; Graves, Brandon, Duesbery, McIntosh, & Pyle, 2011; Lang et al., 2009; Vaughn et al., 2010) or nonexistent (e.g., Fogarty et al., 2014; Hitchcock, Dimino, Kurki, Wilkins, & Gersten, 2011; Kushman, Hanita, & Raphael, 2011; Simmons et al., 2014). These poor outcomes raise the question why carefully designed, research-based interventions have so little an effect on struggling adolescents’ reading skills. As one answer, instruction may be so variable that differences in outcomes between treatment and control groups do not reach statistical significance (Graves et al., 2011; Hitchcock et al., 2011; Lang et al., 2009; Simmons et al., 2014). In other words, teachers’ implementation of the “treatment” condition is often so varied that, in the end, it is difficult to pinpoint exactly what the “treatment” is. Thus, fidelity of implementation –
defined as “the degree to which specified procedures are carried out as planned” - is an essential consideration (Dane & Schneider, 1998, p. 23).

Together, the two studies of this dissertation provide a nuanced description of teacher-level variation in instructional practice, in connection with literacy coaching and student outcomes, including a detailed look at the instructional practices of four teachers implementing their version of a literacy intervention after the formal conclusion of the intervention study. Furthermore, both studies examine mature implementation of the intervention, as all participating teachers were in at least their second year of implementation. This is rare in literacy intervention research, as most of the implementation literature is focused on teachers in their first year of implementation (e.g., Benner, Nelson, Stage, & Ralston, 2011; Burkhauser & Lesaux, 2015; Cantrell, Almasi, Carter, & Rintamaa, 2013; Cantrell et al., 2010; Faggella-Luby & Wardwell, 2011; Graves et al., 2011; Lang et al., 2009; Leko, Roberts, & Pek, 2015; Simmons et al., 2014; Solís, Vaughn, & Scammacca, 2015; Vaughn et al., 2011).

Data for these two studies were gathered within the context of the Strategic Adolescent Reading Intervention (STARI), designed for middle school students scoring two or more years below grade level on standardized reading tests. Reading remediation in the upper grades is challenging for many reasons: first, the literacy skills required for successful reading at this level are complex and varied; second, the amount of instruction needed for struggling readers to “catch up” increases with each of year of schooling; and finally, older struggling readers have often become disengaged from literacy and academics (Biancarosa & Snow, 2004; Lang et al., 2009). Yet STARI has demonstrated a
positive impact on the literacy skills of struggling readers in diverse, high-poverty schools (Kim et al., 2016).

STARI’s theory of change puts forth that building teacher knowledge and capacity and providing engaging topics and materials will improve students’ skills in both basic and deep comprehension. The central mechanism through which this is theorized to occur is structured discussion, or student talk. STARI classes meet for a full class period for the whole school year.

Each STARI lesson plan begins with fluency practice during which students engage in repeated readings of leveled texts, tracking words per minute over time. The topics for these leveled nonfiction passages are related to unit themes in order to build background knowledge and aid with comprehension. For example during the unit on Ninth Ward (Rhodes, 2010), a novel about Hurricane Katrina, fluency passages included information about hurricanes and about the history and culture of New Orleans.

STARI uses the Reciprocal Teaching approach to comprehension instruction, focused on four strategies: clarifying, summarizing, predicting and questioning (Palincsar & Brown, 1984). The formal curriculum instructs teachers to divide their class into two groups, and to meet with one group for a guided reading lesson while the other group engages in partner reading. In guided reading lessons, the teacher previews a section of text that students then read independently, after which the teacher leads discussion of that section (Fountas & Pinnell, 2006). STARI provides direction for each guided reading lesson, including a text introduction with a vocabulary preview and questions for discussion during and after reading. Typically, after reading, STARI assigns students to complete a comprehension activity in their workbooks related to that section of text.
STARI has a number of key features that differentiate it from other adolescent reading interventions. First, STARI is designed around engaging and accessible texts. Each unit is anchored around a novel with an appropriate level of difficulty for middle school students reading at or below the 35th percentile; however, novels were also selected for “characteristics of cognitive challenge: the degree to which readers must work through plot and character ambiguities, resolve diverse perspectives, and use specific background knowledge to bridge gaps in the text” (Kim et al., 2016, p. 10). This combination of appropriate readability levels for struggling students with challenging subject matter was designed to build students’ engagement and inferential comprehension through talk about text. STARI’s second key feature is peer collaboration and student voice. Opportunities for discussion are threaded throughout the curriculum; in addition to partner work and the discussion questions included in every guided reading lesson, each unit includes at least one debate. The third key feature of STARI is integration of lower-level and higher-level reading skills; as described above, lower-level skills like decoding and fluency are connected thematically with the cognitively challenging work of comprehending unit texts.

STARI was introduced to schools in a randomized controlled trial beginning with pilot studies conducted in the 2011-12 school year and concluding with field trials in nine schools between fall 2013 and spring 2015. During this period of time, participating teachers received extensive professional development. Before beginning implementation, all teachers attended a summer institute to receive training in the STARI instructional model. During the school year, teachers received further training through bimonthly meetings in professional learning communities and weekly visits from a coach. Coaches
were expected to make contact with each teacher each week during the school year, but how that time was spent was left to the coach’s discretion. Coaches were required to record in weekly coaching logs the amount of time spent with each teacher, as well as noting how that time was spent (e.g., a meeting, an observation, co-teaching), and what topics the coach and teacher focused on that week (e.g., lesson planning, assessments, guided reading, word study).

Data for the first study of this dissertation (Chapter 2) were collected during the 2013-14 school year. In this study, I used data from observations of the 17 teachers in nine schools, all in the treatment group of the larger RCT, to conduct a nuanced descriptive analysis of fidelity of implementation (FoI). I also analyzed weekly logs completed by three literacy coaches to examine variation in quantity and intensity of coaching. I then compared variation in coaching with variation in FoI, and finally compared FoI to outcomes for students.

After the randomized controlled trial of STARI concluded, one of the nine participating schools sustained implementation of the intervention. Data for the second study (Chapter 3) were gathered in this school during the 2015-16 school year. Through observation and interview of the school’s four STARI teachers, I analyzed the adaptations the teachers made to the curriculum. First, I selected six focal adaptations, three each from two teachers, and evaluated the productivity of each adaptation using criteria from the work of Debarger and colleagues (Debarger et al., 2013). Then, I used an accounts of teaching practice methodology (Simon & Tzur, 1999) to define an orientation toward the curriculum for each of the four teachers (Remillard & Bryans, 2004). Finally, I compared each teacher’s orientation toward the curriculum with
Coburn’s (2004) framework describing teachers’ responses to policy initiatives, to determine whether each teacher displayed assimilation or accommodation to intervention principles.

In the final chapter, Chapter 4, I discuss the connections among the findings of the two studies, as well as the studies’ implications for practice and research.

This dissertation offers a number of important contributions to the research on literacy interventions, fidelity of implementation and curriculum studies. First, my research provides a nuanced description of fidelity of implementation that is rare within literacy intervention research. Secondly, unlike much previous research, my studies analyze data on coaching, instruction, and student outcomes, allowing connections to be drawn from coaching to instruction and from instruction to student outcomes. Third, both studies investigate mature implementation of an intervention, rather than focusing on teachers in their first year of implementing a new instructional program. Fourth, the second study of the dissertation presents a detailed picture of successful, sustainable implementation of an intervention with a focus on adaptations made by experienced teachers.

A better understanding of variation in implementation is critical for future research in literacy interventions, as teacher-level adaptations may make interventions either more or less effective than the curriculum as written. The findings of this dissertation may allow curriculum developers to provide more tailored professional development to support teachers’ implementation of literacy intervention curricula.
Chapter 2.

Fidelity of Implementation of an Adolescent Reading Intervention

Reviews of the research on program implementation have found that no intervention has demonstrated 100% fidelity, and few attain greater than 80% (Durlak & DuPre, 2008). In educational interventions, teacher-level implementation differences of 20-30% are common (Durlak, 2010). Yet effect sizes are typically two to three times larger for providers with higher levels of adherence to the intended intervention (Durlak & DuPre, 2008). In order to improve and replicate existing interventions, it is as important to document the process of implementation as it is to measure outcomes (Meyer, Miller, & Herman, 1993).

A number of literacy intervention studies have employed literacy coaches to provide one-on-one professional development to teachers, in an effort both to standardize and to improve teachers’ implementation of the intervention. Yet significant research has demonstrated vast variation in the practices of literacy coaches (e.g., Atteberry & Bryk, 2011; Deussen, Coskie, Robinson, & Autio, 2007; Ippolito, 2010; Scott, Cortina, & Carlisle, 2012). Perhaps because of this variation, findings on the effectiveness of literacy coaching are mixed: some research (e.g., Biancarosa, Bryk, & Dexter, 2010; Sailors & Price, 2015) has demonstrated significant impacts of coaching on instruction and/or student outcomes, while other studies (e.g., Garet et al., 2008; Matsumura, Garnier, Correnti, Junker, & Bickel, 2010) have found no relationship between coaching and outcomes for students or teachers. Therefore, further research into the practices of literacy coaches, and the relationships among coaching, instruction and student achievement is warranted.
The present study therefore seeks to document the process of implementation of an adolescent literacy intervention (STARI), and to better understand sources of variation in implementation, by providing a nuanced description of teacher-level variation in instructional practice, in connection with literacy coaching and student outcomes.

**Conceptual Background**

Dane and Schneider (1998) propose a five-part framework to describe and operationalize fidelity of implementation (FoI). The first aspect is *adherence*, meaning compliance or faithful replication of the intervention as designed. Second, *dosage* refers to the quantity of intervention to which participants are exposed. Third, *quality of delivery* describes the skill with which the various components of the intervention are executed. Fourth, *participant responsiveness* measures the degree of engagement with the intervention demonstrated by participants. Finally, *program differentiation* describes the extent to which the particular intervention can be distinguished from other interventions.

**Fidelity of Implementation and Reading Interventions**

Typically, research on adolescent reading interventions reports an overall FoI score, usually operationalized as adherence (e.g., Denton, Wexler, Vaughn, & Bryan, 2008; Diliberto, Beattie, Flowers, & Algozzine, 2008; Faggella-Luby & Wardwell, 2011; Manset-Williamson & Nelson, 2005; Pyle & Vaughn, 2012; Vaughn et al., 2010; Vaughn et al., 2012; Wexler, Vaughn, Roberts, & Denton, 2010). This score is generally reported in the form of an average across teachers to support the claim that overall, the intervention was implemented with high fidelity. Meanwhile, relationships between FoI and student outcomes are rarely explored. For example, one study of the popular adolescent reading intervention program READ 180 (Houghton Mifflin Harcourt, 2015)
conducted three observations per teacher, conceptualizing FoI as adherence and rating teachers on a three-point scale for each of seven instructional activities (Kim, Capotosto, Hartry, & Fitzgerald, 2011). However, they did not consider FoI in their analysis of student outcomes. As an additional example, in a study of another research-based adolescent reading intervention, Collaborative Strategic Reading (CSR), researchers conducted four observations of each teacher, rating teachers on adherence to each curriculum component as well as global instructional quality (Vaughn et al., 2011). However, the researchers did not consider teacher-level variation in implementation, nor analyze the relationship of FoI to student outcomes. Furthermore, none of the research cited above considers variation across aspects of fidelity, for instance whether some dimensions of fidelity may be easier to achieve than others, or whether some may have a greater influence on student outcomes. In an attempt to fill this gap, the present study offers a detailed descriptive analysis of adherence, dosage, quality and participant responsiveness throughout the implementation of an adolescent reading intervention.

To date, only a few reading intervention studies have taken this nuanced view of fidelity of implementation. For example, in Fogarty and colleagues’ (2014) study of an intervention implemented by 14 middle school teachers, the researchers created an FoI protocol based on Dane and Schneider’s (1998) five components, measuring adherence, quality, participant responsiveness and program differentiation through observation, and dosage through teacher self-report. A confirmatory factor analysis revealed that these five components loaded onto one factor. Although the researchers found no main effects of their intervention, they found a significant relationship between the FoI factor and student achievement on both standardized and researcher-designed comprehension measures,
such that students in high fidelity classrooms showed more than three times the growth of students in low fidelity conditions.

In addition, in an adolescent reading intervention involving 21 teachers, researchers used a five-item observation protocol to measure adherence and quality (Benner et al., 2011). They found that overall fidelity explained 22% of the variance in gains in basic reading and 18% of student gains in passage comprehension. Specifically, the single adherence item and one quality item ("reteaches when needed") contributed to student gains; the other three quality items made no statistically significant contribution.

Finally, in a summer reading intervention for struggling first and second graders, Zvoch (2012) measured the adherence of the 24 participating teachers through observation, and measured participant responsiveness through teacher ratings of student engagement. He found that 51% of the variability in students’ oral reading fluency growth was due to between-group difference. Furthermore, while adherence had no direct association with student outcomes, adherence moderated the relationship between participant responsiveness and growth, such that engaged students grew more in classrooms where teachers adhered closely to the intervention as written. Collectively, these three studies indicate the importance of FoI to student outcomes in reading interventions, as well as the complexity of measuring FoI in a valid and meaningful way. None of these studies analyzed predictors of FoI or considered the role that professional development may play in variation in implementation, both of which are addressed in the present study.

Implementation and Professional Development
In order to improve teachers’ practice both within and outside of evidence-based programs, a consensus exists that high-quality professional development should be sustained, focused and intensive, rather than consisting of one-shot, disconnected workshops as in years past (Correnti, 2007; Scher & O'Reilly, 2009; Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). These recommendations have led many schools, districts, and intervention programs to employ literacy coaches to provide on-site, sustained, individualized professional development. The guiding theory of change puts forth that coaching will improve teachers’ instructional practice, which will lead to improvement in student outcomes (Scher & O'Reilly, 2009; Yoon et al., 2007).

However, the few randomized controlled trials that have assessed literacy coaching as a specific aspect of an intervention demonstrate mixed outcomes. In one such study, Garet and colleagues (2008) compared second grade teachers who were coached and attended a training institute with those who only attended the institute, as well as a control group. Although both treatment groups demonstrated a higher level of knowledge and implemented more explicit instruction than the control group, there was no difference between the coached and uncoached groups on any measure, and neither treatment showed any effect on student outcomes. In a randomized controlled trial of literacy coaching for fourth and fifth grade teachers, treatment teachers self-reported increasing target instructional practices, but the study demonstrated no effect on student outcomes (Matsumura et al., 2010). However, the student outcome in both studies - scores on the state standardized test - may simply be too far downstream to show measurable improvement as a result of teacher coaching, and no interim data was collected on students’ skills.
In contrast, Biancarosa, Bryk and Dexter’s (2010) analysis of the Literacy Collaborative coaching model in kindergarten through second grade demonstrated that coaching was associated with student achievement. Likewise, Sailors and Price (2015) conducted a randomized controlled trial of the SIPIC coaching model for second through eighth grade teachers. In this study, all teachers attended a two-day professional development workshop, and treatment teachers received coaching in addition. Classroom observations revealed that coached teachers implemented more target instructional practices than control teachers, and their students made greater gains on a standardized assessment of reading comprehension (i.e., the GRADE (Pearson Education, 2017)).

Although not specific to literacy, Kraft and Blazar (2016) found evidence that some of this variation in coaching outcomes may be due to program features. These researchers conducted a two-year randomized controlled trial of the MATCH Teacher Coaching program for teachers of all grade levels. In their first cohort, coached teachers’ scores on an index of instructional practices showed an increment equivalent to more than the average difference between beginning and experienced teachers. In their second cohort, just one year later, the coaching program demonstrated no results. Between the two years, the MATCH program had overhauled its model, increasing the number of teachers assigned to each coach, decreasing the amount of coaching each teacher received, and standardizing the content of coaching sessions. Thus, the authors suggested that structural factors like these may influence the effectiveness of coaching.

Similarly, in their study of primary-grade coaches in Reading First schools, Bean and colleagues used a median split to categorize schools based on the amount of time coaches spent working with teachers; they found that a significantly higher proportion of
students scored proficient on reading assessments in schools where coaches devoted more time to teachers (Bean, Draper, Hall, Vandermolen, & Zigmond, 2010). Other descriptive analyses of literacy coaching (e.g., Atteberry & Bryk, 2011; Deussen et al., 2007; Ippolito, 2010; Scott et al., 2012) further illustrate the vast variability in what literacy coaches do. For instance, in Sailors and Price’s study (2015), coaches visited teachers between one and sixteen times per year, with total coaching time per teacher ranging from 45 to 830 minutes; in Garet et al.’s (2008) study, total coaching time over the course of the school year ranged from one to 173 hours per teacher. The literature also reports significant variation in how coaches spend that time; for example, Deussen and colleagues reported some coaches spending over 90% of their time working directly with teachers, while others spent almost half their time managing paperwork, data and assessments.

Few studies to date have attempted to link variation in coaching with variation in instructional practice. Garet and colleagues (2008) examined predictors of variation in coaching time, and found that the only predictors were teaching credential, such that teachers who lacked a regular credential received more coaching time, and logistical factors, such as the total number of teachers assigned to a coach. They did not examine predictors of how coaches spent their time, nor did they link coaching time to teacher or student outcomes. Sailors and Price (2015) attempted to link coaching with instruction, determining the effect sizes of five specific coaching practices. However, the effect sizes were all similar in magnitude (.64-.74), with overlapping confidence intervals, providing little information about the relative effectiveness of each practice. Since previous research demonstrates tremendous variation in coaching practices, which likely plays a
role in the effectiveness of coaching, it is essential to closely examine the process of coaching in connection with instruction and student outcomes.

The Present Study

The present study provides a nuanced description of teacher-level variation in fidelity of implementation, variation in the coaching these teachers received, and the relationship between variation in instruction and proximal student outcomes. The study is framed around the research questions:

RQ1: How does implementation of a research-based reading intervention vary at the teacher level?
RQ2: How do literacy coaches’ interactions with teachers vary?
RQ3: Is variation in coaching associated with variations in instructional practice?
RQ4: What is the relationship between fidelity of implementation and student outcomes?

The present study offers a number of important contributions to the literature. First, as described above, studies of adolescent reading interventions usually report an overall adherence score, averaged across teachers, with no further exploration (e.g., Denton et al., 2008; Diliberto et al., 2008; Faggella-Luby & Wardwell, 2011; Manset-Williamson & Nelson, 2005; Pyle & Vaughn, 2012; Vaughn et al., 2010; Vaughn et al., 2012; Wexler et al., 2010). The present study provides a more complete and nuanced look at FoI than is typical in reading intervention research. In order to improve implementation of interventions, it is essential to know both where teachers struggle and which aspects of FoI may be most important in affecting student outcomes.

Secondly, unlike much existing research, the present study analyzes sufficient data to address all aspects of the theory of change by linking coaching to teacher practice and teacher practice to student outcomes. Some existing studies analyze the effects of coaching on student outcomes without considering the mediating factor of changing
instructional practices (e.g., Biancarosa et al., 2010; Elish-Piper & L’Allier, 2011; Lovett et al., 2008; Marsh et al., 2008), while others consider associations between coaching and instructional practices with no direct evidence that these changes to instruction have a positive effect on student achievement (e.g., Brown et al., 2007; Cantrell & Callaway, 2008; Coburn & Woulfin, 2012). The present study attempts to isolate, albeit in a nonexperimental way, the specific role of coaching in relation to instruction and student outcomes.

Third, the vast majority of intervention research is short-term, collecting data in the first year of implementation (e.g., Benner et al., 2011; Cantrell et al., 2013; Cantrell et al., 2010; Faggella-Luby & Wardwell, 2011; Graves et al., 2011; Lang et al., 2009; Simmons et al., 2014; Solís et al., 2015; Vaughn et al., 2011). Therefore, these studies examine teachers’ practice while teachers are still in the process of learning a new curriculum or instructional approach. Typically, student achievement increases as teachers gain experience in implementing an intervention (Biancarosa et al., 2010; Lovett et al., 2008). All teachers in the present study were in at least their second year of implementation, providing the most valid context for research, as “complicated interventions require more time to be conducted properly…and evaluations conducted before implementation is sufficiently established will not do justice to the true impact of the intervention” (Durlak & DuPre, 2008, p. 343). Thus, this study presents a picture of FoI during a more mature stage of program implementation than is typical.

Fourth, much of the literature on coaching and instruction (e.g., Bean et al., 2010; Brown et al., 2007; Lovett et al., 2008; Marsh et al., 2008; Marsh, McCombs, & Martorell, 2012; Matsumura et al., 2010; Stephens et al., 2011) relies on teachers’ self-
report of changes in beliefs and instructional practices and coaches’ retrospective self-report through surveys and interviews. The present study is grounded in direct observations of teachers and detailed weekly coaching logs spanning an entire school year, enhancing validity of measures and offering an important contribution to this research base.

Finally, much existing research compares literacy coaching to student outcomes such as standardized tests (e.g., Biancarosa et al., 2010; Marsh et al., 2012), which may be too distal to show results. It is difficult for adolescent literacy interventions to achieve large effects on standardized assessments (Scammacca, Roberts, Vaughn, & Stuebing, 2013). Because literacy coaching is hypothesized to directly affect instruction, which then affects student learning, some research suggests that this complex process may result in effects on standardized student outcomes that are too small to detect (e.g., Garet et al., 2008). Therefore, I hypothesize that coaching may have a stronger relationship with extremely proximal measures of student achievement, which may be more directly tied to teacher practice and easier to affect within the brief period of an intervention. Thus, I focused on the relationship between coaching and students’ workbook completion and daily oral reading fluency, in the hopes of observing associations between coaching and proximal outcomes which may be suggestive of relationships with longer-term outcomes.

**Methods**

**Participants**

Participants include three literacy coaches, 17 sixth, seventh and eighth grade teachers in nine schools in four districts in Massachusetts, and the 287 students they collectively taught, aged 11-15. All participants in the present study were assigned to the
treatment condition of the larger randomized controlled trial of STARI. STARI classes met for a full class period for the whole school year. Individual schools’ different schedules meant that the configuration of STARI classes varied at the school level, with class periods lasting between 45 and 60 minutes, and meeting between three and five times per week. In all cases, STARI classes met for the same frequency and duration as students’ other core academic classes.

Certified teachers at the nine participating schools were recruited to teach STARI. All participating teachers provided written informed consent, entailing their willingness to participate in professional development, implement the curriculum, collect and return student data, and participate in FoI observations. Teachers’ experience ranged from 6 to 35 years ($M=15.44$, $SD=7.61$). All but one had attained a Master’s degree, and two had earned doctorates (one Ed.D. and one J.D.). Fourteen were female and three male, and the majority were White, with two African-Americans, one Latina and one Native American. All were certified in subject areas including reading, English and special education.

**Professional development and teacher support.** Prior to beginning implementation, all teachers attended a summer institute to receive training in the STARI instructional model. During the school year, teachers received further training through bimonthly meetings in professional learning communities and weekly visits from a coach. The STARI project hired three experienced literacy coaches, each of whom served multiple participating schools (see Figure 1), providing support in STARI implementation in addition to any school-based coaching teachers already received. Coaches were hired at different times as the project grew to scale, and their training was largely informal and conducted by current team members. The three coaches met weekly
to collaborate around coaching strategies. Coaches were expected to make contact with each teacher each week during the school year, but how that time was spent was left to the coach’s discretion.

**Measures**

**Observations.** Each teacher was observed by trained research assistants at least twice, at the beginning and end of the school year, for a full class period. Observations were scheduled in advance, and teachers were aware that the observer’s purpose was to assess FoI. Principal investigators designed a 44-item observation protocol with selected features of instruction (e.g. *Teacher sets a purpose for reading the next section of the novel; Teacher directs students to silently read particular text chunks and then stop for discussion*) that reflected the core components of STARI. For each item, teachers were rated on a 0-2 scale, operationalized slightly differently for different instructional practices (e.g., 2=present, 1=partly present, 0= not present, or 2= continually, 1=occasionally, 0=not present). In addition, observers noted the number of minutes spent on each component of the lesson (i.e., fluency practice, comprehension instruction). Seventeen percent of observations were double-coded, and interrater reliability was moderate to high ($K = .84$).

In order to allow a nuanced description of FoI, each item on the observation protocol was coded using Dane and Schneider’s (1998) framework for aspects of fidelity. The author and the designer of the intervention curriculum engaged in post hoc collaborative coding to determine which aspect of fidelity was best represented by each instructional practice, discussing each protocol item until agreement was reached (Smagorinsky, 2008). This coding process resulted in a 17-item adherence scale,
including such items as *Teacher poses literal ("right there") questions*; a 9-item scale measuring quality, including such items as *Teacher asks follow-up questions to elicit fuller or clearer student responses*; and a 6-item scale measuring participant responsiveness, including such items as *Students ask each other and answer follow-up questions or comment to partner*. Twelve items from the observation protocol, mostly pertaining to classroom management and the assignment of homework, were judged not to represent any of Dane and Schneider’s components of fidelity, and were therefore dropped from these analyses (see Appendix A for original observation protocol as well as items for each scale). Initial analyses suggest that the protocol was a reliable tool; Cronbach’s alpha for the full scale was .87, which is considered excellent, and alphas for the three subscales ranged from .73 to .78, considered adequate to very good (Kline, 2011). Furthermore, adherence, quality and participant responsiveness were all strongly inter-correlated (correlations ranged from .68 to .86, all ps < .05).

The total number of times each teacher was observed varied substantially, from two to twelve times, for a variety of logistical reasons (e.g., if teachers taught more than one section of STARI, all sections were observed; some teachers had shorter class periods and divided lesson components across instructional days). Therefore, to ensure a similar comparison across teachers, I chose to focus on two complete observations per teacher. For each teacher, I selected the first time he or she was observed teaching each lesson component as *Wave 1* (W1), and the last time he or she was observed teaching each component as *Wave 2* (W2). For some teachers, each observation took place in a single day; for others, each spanned multiple class periods. W1 observations took place
between December and March, and W2 observations took place in May and June. A minimum of three months elapsed between each teacher’s W1 and W2 observation.

Control students were scattered across numerous alternative classes in the “business-as-usual” condition, making observations of control classrooms impossible. Therefore, we were unable to assess program differentiation. Of the few existing studies taking a multi-component approach to understanding FoI, only Fogarty et al (2014) assessed program differentiation, suggesting that this aspect is uniquely challenging to assess.

**Coaching logs.** Over the course of the school year, the three STARI coaches completed detailed weekly logs in which they recorded the type of contact they had with each teacher (ranging from an email check-in to co-teaching or demonstration teaching); the quantity of contact; and the focus of coaching in both process (e.g., lesson planning, routines, management) and content (e.g., fluency, guided reading, partner work).

**Student outcomes.**

**Workbooks.** Daily assignments for each unit are organized into workbooks, where students practice comprehension skills and respond to questions about assigned texts. Workbooks require students to engage in complex literacy skills; for example, the sample workbook page in Figure 2 requires students to work with a partner to make an inference about how the main character, Lanesha, is feeling; to infer that another main character has died, which is not directly stated in the text, and then to make a prediction about what Lanesha will do. Thus, successful workbook completion indicates students’ high-level comprehension of text. Workbooks were collected and coded for completion. A page was coded as “completed” if a student had made some attempt at meaningful engagement
with the content. The number of completed workbook pages ranged from 0 to 290 pages ($M=164.77$, $SD=61.16$). This wide range suggests that the number of workbook pages a student completed reflects meaningful differences in students’ interaction with the curriculum.

**Daily oral reading fluency.** Fluency workbooks, in which students recorded their daily number of words per minute (WPM) were also collected. Students’ growth in oral reading fluency was determined by comparing WPM at the beginning and end of the school year. In addition, daily WPM were summed to calculate the total number of words each student read during fluency practice over the course of the school year.

**Data Analyses**

To address RQ1, “How does implementation of a research-based reading intervention vary at the classroom level?,” I first used teacher observation data to conduct a descriptive analysis of FoI. Next, I examined predictors of FoI (e.g., years of teaching experience, highest degree earned). The teachers in this study were nested within schools (see Figure 1), and since it is assumed that teachers in the same school may share many common, unobserved experiences that could affect FoI, multilevel modeling would be the ideal method of data analysis. However, imprecision caused by the small sample size would limit interpretability of multilevel models. Therefore, I used OLS regression to examine relationships between FoI and other variables. I also conducted exploratory analyses using multilevel models to partition the variance in FoI between the teacher and school levels; however, these results should be interpreted with caution.

To address RQ2, “How do literacy coaches’ interactions with teachers vary?,” I began by conducting a descriptive analysis of the practices in which the three coaches
engaged. I then coded the level of intensity of coaching practices according to Bean’s (2004) coding scheme. According to this scheme, level 1 indicates low-intensity coaching practices, such as an email or a quick check-in; level 2 indicates moderate-intensity coaching practices, such as co-planning lessons or leading professional development presentations; and level 3 indicates high-intensity coaching practices, such as co-teaching or observing and providing feedback.

To address RQ3, “Is variation in coaching associated with variation in instructional practice?,” I examined predictors of coaching time and intensity, as well as analyzing coaching as a predictor of W2 FoI. To examine predictors of coaching time and intensity, I created a series of regression models, beginning with teacher-level demographic predictors such as years of experience and highest degree earned. I next examined W1 FoI scores as predictors of coaching time and intensity; in other words, analyzing whether a teacher’s skill in implementing the intervention near the beginning of the school year predicted how, and how much, a coach spent time with a teacher. Then I investigated whether coaching time or intensity predicted W2 FoI, controlling for W1 FoI; in other words, I examined whether teachers who had received more coaching and/or more intense coaching showed greater improvements in FoI than other participants. In addition, I conducted exploratory analyses using multilevel models in order to partition variation in coaching time and intensity across the teacher and coach levels.

Finally, to address RQ4, “What is the relationship between FoI and student outcomes?,” I created a series of multi-level models with students nested by teacher, allowing the relationship between FoI and student outcome to vary randomly at the

---

1 Due to the small sample size, I examined demographic predictors and FoI predictors in separate models.
student level. I examined whether a teacher’s FoI predicted students’ workbook completion or daily oral reading fluency.

Results

RQ1: Variation in Fidelity of Implementation

In order to better understand teacher-level variation in FoI, I first conducted a descriptive analysis of teacher observation data, and then examined teacher-level predictors of FoI at W1. As in similar large-scale, long-term studies of complex adolescent reading interventions (e.g., Cantrell et al., 2014; Feldman, Feighan, Kirtcheva, & Heeren, 2012; Simmons et al., 2014) overall fidelity of implementation was moderate to high. On average, teachers implemented 78% (25 of 32; SD=5.36) of the instructional practices on the observation tool at W1, and 79% (25.41 of 32; SD=4.93) at W2. However, this average conceals significant variation, from a low of 9 instructional practices (28%) to a high of 31 (97%). Therefore, my next step was to explore that variation.

I began by examining variation in implementation across the theoretical components of FoI for which the observation tool was coded: adherence, quality and participant responsiveness (see Table 1 for complete descriptive statistics). On average, at W1, teachers implemented 6.47 of the 7 practices on the fluency adherence scale (Min=5, Max=7, SD=0.62), 7.44 of the 10 practices on the comprehension adherence scale (Min=3, Max=10, SD=2.19), 6.17 of the 9 items on the quality scale (Min=1, Max=9, SD=2.13), and 4.17 of the 6 items on the participant responsiveness scale (Min=1, Max=6, SD=1.42). Item-level scores were averaged to create a mean score for each teacher for each component of the observation scale. Paired t-tests revealed that at W1,
adherence was statistically significantly higher than quality, $t(16)=3.35$, $p=.004$, and quality was statistically significantly higher than participant responsiveness, $t(16)=2.08$, $p=.05$. Furthermore, adherence to the fluency portion of the lesson was statistically significantly higher than adherence to the comprehension portion, $t(16)=2.66$, $p=0.02$.

Next, I examined variation in FoI across the two time points, also shown in Table 1, and found that there were no statistically significant differences in W1 and W2 scores on overall FoI, adherence, quality or participant responsiveness. W2 FoI scores displayed a similar pattern to W1, with fluency adherence statistically significantly higher than comprehension adherence, and both adherence and quality statistically significantly higher than participant responsiveness. There were differences between waves, however, in dosage, or the amount of time devoted to each portion of the lesson. At W1, teachers spent an average of 19.59 minutes on fluency instruction ($SD=4.20$), and 26.59 minutes on comprehension ($SD=17.54$), with no statistically significant difference between the two, $t(16)=-1.63$, $p=0.12$. However, at W2, the amount of time spent on fluency instruction ($M=14.88$ minutes, $SD=5.16$) was statistically significantly lower than the time spent on comprehension ($M=34.29$ minutes, $SD=14.49$, $t(16)=-5.20$, $p<.001$), as well as being lower than the time spent on fluency at W1, $t(16)=-4.18$, $p<.001$. There was no statistically significant difference in the amount of time spent on comprehension instruction between the two waves of observation, $t(16)=-1.67$, $p=0.12$. Thus, on average, between the two observations, teachers shifted from spending a relatively equal amount of time on fluency and comprehension instruction, to spending less time on fluency than on comprehension. As the STARI curriculum was designed to allocate more time to
comprehension than to fluency, this shift reflected an increase in adherence to the suggested pacing of daily lesson plans between W1 and W2.

Finally, I created a multilevel model, with teachers nested by school, to conduct an exploratory analysis of the proportion of variation at the teacher vs. the school level. This analysis suggested that the majority of the variation in FoI occurred at the teacher level, with only about 10% of the variation occurring at the school level.

**Predictors of variation in fidelity of implementation at W1.** I next examined predictors of teachers’ FoI at W1, including teacher demographic characteristics such as highest degree earned, whether or not the teacher held a certification in the teaching of reading, years of teaching experience, and years of experience teaching STARI specifically (see Table 2). Holding a reading certificate was the only predictor statistically significantly associated with FoI, such that on average, teachers with a reading certificate scored 14.73 points higher on the 64-point FoI scale ($p=.02$), controlling for all other predictors.

**RQ2: Variation in Coaching**

In order to better understand literacy coaches’ interactions with teachers, I conducted a descriptive analysis of detailed weekly logs completed by the three STARI coaches, reporting a total of 407 contacts with teachers over the course of the school year. The total amount of time a coach and a teacher spent working together during the school year varied widely among teachers, ranging from a total of 13.83 hours to 38.92 hours, with a mean of 25.62 hours ($SD=7.89$) (see Table 3 for complete descriptive statistics). There were no statistically significant differences between coaches in their average amount of coaching time per teacher.
Also reported in Table 3, I coded coaches’ contact with teachers by level of intensity, according to Bean’s (2004) 3-point coding scheme, described above, with level 1 indicating low-intensity coaching practices and level 3 indicating high-intensity coaching practices. Since most coach-teacher interactions involved more than one coaching practice (e.g., an email check-in and an observation), I coded each interaction with the highest-level activity reported. Across the school year, nearly half of coaching interactions (46%) included high-intensity activities, 22% included medium-intensity activities, and 32% included only low-intensity activities. Coach 2’s average level of intensity (\(M=2.45, SD=.07\)) was statistically significantly higher than that of Coach 1 (\(M=1.95, SD=.08, t(227)=-4.78, p<.001\)) or Coach 3 (\(M=2.04, SD=.07, t(295)=4.15, p<.001\)). Both Coach 2 and Coach 3 engaged in more high-intensity practices (i.e., lesson observations and co-teaching) than Coach 1 did; however Coach 3 also engaged in more low-intensity practices (i.e., meetings). Therefore, of the three coaches, Coach 2 reported the most frequent engagement in high-intensity coaching practices (64% of interactions), compared with 44% and 29% for Coaches 3 and 1, respectively. Variation in coaching does not seem to be explained by differences in the student and teacher populations that each coach served. These populations appeared similar at baseline, as there were no statistically significant coach-level differences in students’ fall WPM, \(F(2, 278)=1.59, p=0.21\), or teachers’ W1 FoI scores, \(F(2, 14)=0.99, p=0.40\).

Exploratory analyses using multilevel models, with teachers nested within coaches, suggested that far more variation occurred at the teacher level than at the coach level. For coaching time in particular, less than 1% of the variation occurred between coaches. More of the variation in coaching intensity occurred at the coach level.
(approximately 25%), although again a majority of the variation occurred at the teacher level.

**RQ3: Variation in Coaching and Variation in Instructional Practice**

**Predictors of variation in coaching.** To understand the associations between coaching and instructional practice, I analyzed predictors of variation in coaching (see Table 2). First, I examined whether teacher characteristics, such as certification or years of experience, predicted the amount of time a coach would spend with that teacher. No demographic predictors were statistically significantly associated with coaching time. Next, I examined whether W1 FoI scores would predict coaching time, and found that fluency adherence at W1 predicted coaching time such that scoring one point lower on the 14-point scale was associated with, on average, receiving 2.73 more hours of coaching time over the course of the school year (\(SE=1.10, p=.03\)), controlling for all other aspects of FoI. Neither demographic predictors nor W1 FoI scores predicted the intensity with which teachers were coached.

**Variation in coaching as a predictor of fidelity of implementation.** Also reported in Table 2, I explored whether coaching time or intensity would predict FoI at W2 when controlling for FoI at W1. In other words, I wondered whether teachers who had received more coaching and/or more intense coaching would show greater improvements in FoI than other participants. In this model, W1 FoI was not a statistically significant predictor of W2 FoI (\(\beta=0.42, SE=0.22, p=.07\)). When controlling for FoI at W1, there was a small, statistically significant negative association between coaching time and W2 FoI, such that one additional hour of coaching was associated with a 0.6-point decrement on the 64-point FoI scale (\(SE=0.31, p=.05\)). There was no statistically
significant association between intensity of coaching and W2 FoI when controlling for W1 FoI ($\beta=-7.56$, $SE=7.63$, $p=.34$).

**RQ4: Fidelity of Implementation and Student Outcomes**

Finally, I examined FoI in relation to two proximal student outcome measures: the number of pages a student completed in the comprehension workbook and a student’s daily oral reading fluency. For this analysis, I created a series of multilevel models, with students nested by teacher. Since there were no statistically significant differences in W1 and W2 FoI, average FoI scores were used for each teacher. I first ran a model examining overall FoI as a predictor of each outcome, and then ran separate models using each FoI subscale (adherence, quality, participant responsiveness and dosage) as predictors (see Table 4 for results). FoI predicted workbook completion, such that scoring one point higher on the 64-point FoI scale was associated with students completing 2.49 more workbook pages on average ($SE=2.38$, $p=.05$). The relationship was particularly strong between adherence and workbook completion, such that scoring one point higher on the 34-point adherence scale was associated with students completing 5.69 more workbook pages on average ($SE=2.38$, $p=.02$). Quality, participant responsiveness and dosage did not predict workbook completion.

There was no relationship between any aspect of a teacher’s FoI score and his or her students’ growth in oral reading fluency; however, all aspects of FoI except dosage predicted the total number of words a student read over the course of the school year. For instance, a teacher who scored one point higher on the FoI scale had students who read, on average, 121 more words during fluency practice over the course of the school year ($p=.003$). In other words, a 2-point increment on the 64-point FoI scale was equivalent to
a student reading approximately one additional passage. Since the mean number of passages read was 21.44 ($SD=6.37$), a 2-point increment on the FoI scale was associated with approximately a 5% increment in a student’s total reading during fluency instruction over the course of the school year.

**Discussion**

In this study, I used data from observations of the 17 teachers in nine schools to conduct a nuanced descriptive analysis of fidelity of implementation. I also analyzed weekly logs completed by three literacy coaches to examine variation in quantity and intensity of coaching. I then compared variation in coaching with variation in FoI, and finally compared FoI to outcomes for students. This research provides a nuanced description of fidelity of implementation that is rare within literacy intervention research, including drawing connections among coaching, instruction and student outcomes. This study suggests three key findings: first, substantial teacher-level variation occurred, despite the teachers’ experience and the substantial professional development they received, and holding a certificate in the teaching of reading was the only predictor of higher levels of FoI. Secondly, coaching seemed to be mostly unrelated to fidelity of implementation, and finally, FoI was associated with student literacy outcomes.

**Variation in Implementation**

The findings of the present study make a number of unique contributions to the literature on literacy coaching and adolescent reading interventions. First, this study provides a more complete and nuanced look at FoI than is typical in reading intervention research. The teachers in this study performed with a level of fidelity that is comparable to other research; however, the average masked tremendous variation. It is unrealistic to
expect that the students of a teacher who implemented 9 of 32 target instructional practices would demonstrate similar outcomes to the students of a teacher who implemented 31 of 32 practices. Thus, it seems plausible that, as other researchers suggest (Graves et al., 2011; Hitchcock et al., 2011; Lang et al., 2009; Simmons et al., 2014), overall low effect sizes of adolescent reading interventions may be a result of variation in implementation. In other words, skilled implementers of an intervention may achieve excellent results in their classrooms, while more struggling teachers may achieve poorer results, leading to overall small or nonsignificant effects across an intervention sample. This finding points to the importance of a more in-depth examination of fidelity of implementation in future reading intervention research.

However, knowing that teachers’ implementation varies is of minimal utility without an understanding of why teachers’ implementation varies. Although the sample was too small to draw definitive conclusions, exploratory analyses suggest that the vast majority of variation in the present study occurred at the teacher level, rather than the school level; in other words, teachers within a school were more different from one another than they were, as a group, different from teachers at another school. This suggests that variation in implementation cannot be explained by differences in school schedules, culture or administrative support. Nor can they be explained by two common teacher-level proxies for quality: years of teaching experience or highest degree earned.

How, then, might this variation be explained? First, it seems clear that some aspects of implementation were easier than others. Teachers consistently scored higher on measures of adherence than on measures of quality or participant responsiveness. Furthermore, teachers overall did a better job adhering to the relatively straightforward
routines of the fluency portion of the lesson plan than to the more complex activities of comprehension instruction, such as leading discussion in guided reading lessons.

Secondly, although all teachers held appropriate subject area certifications, teachers who held a certificate in the teaching of reading earned statistically significantly higher FoI scores than teachers without this certificate. This suggests that despite the intensive supports teachers received – detailed daily lesson plans, summer training and weekly coaching – prior expertise in the teaching of reading played a role in successful implementation of the intervention. Some research suggests that teachers need to understand the theory and deep principles underlying an intervention in order to implement it successfully (Han & Weiss, 2005; McLaughlin & Mitra, 2001; Penuel, Phillips, & Harris, 2014); it seems likely that a reading certificate better equipped teachers to understand those principles.

Variation in Coaching

The present study demonstrates that even within the context of a single intervention, substantial variation exists in the amount and intensity of coaching that teachers receive. This is in line with previous work demonstrating the tremendous variety of practices performed by literacy coaches (e.g., Atteberry & Bryk, 2011; Bean et al., 2010; Deussenn et al., 2007; Ippolito, 2010; Scott et al., 2012). However, importantly, most previous work has assumed that this variation occurs at the coach level (e.g., Bean et al., 2010; Camburn, Kimball, & Lowenhaupt, 2008; Marsh et al., 2008), with one study even going so far as to define five categories of coaches based on their interactions with teachers (Deussen et al., 2007). The present study, on the other hand, suggests that the majority of variation occurs at the teacher, rather than the coach, level. Although the
amount of time a teacher spent working with a coach in this study ranged from a minimum of 14 hours across the school year, to a maximum of more than triple that, exploratory analyses suggest that almost none of the variation in coaching time occurred at the coach level. Perhaps, then, the variation that has been demonstrated in previous research reflects, not differences in coaches, but differences in coaches’ assessments of teachers’ need and desire for coaching.

While the amount of time coaches spent working with teachers appeared to be explained almost entirely by teacher-level variation rather than coach-level variation, the way coaches spent that time appeared to be somewhat more dependent on the coach. Coaches 2 and 3 engaged in significantly more high-intensity coaching practices than Coach 1 did (i.e., observations and co-teaching); however Coach 3 also engaged in significantly more low- and moderate-intensity coaching practices (i.e., meetings). Therefore, Coach 2’s interactions with teachers were, on average, more intense than those of Coach 1 or Coach 3.

As with variation in FoI, however, knowing that coaching practices vary is less important than understanding why they vary. If, as the data suggest, each coach adapted her coaching to meet the needs of individual teachers, to what information did coaches attend in determining those needs? Presumably, these decisions were based largely on coaches’ personal interactions with teachers, and the data in the present study provide an incomplete picture. However, findings did reveal that teachers who demonstrated lower levels of adherence to STARI’s fluency routines at W1 received more coaching time, on average. Thus, it appears that struggles implementing the fluency routine may have been the primary indicator coaches used to identify teachers in need of help. Since adherence
to the fluency routine was the area on the observation tool where overall average scores were highest (as compared to comprehension adherence, quality or participant responsiveness), it seems logical that coaches would have identified the relatively few teachers who struggled in this area as the teachers most in need of coaching support.

**Coaching, Implementation and Outcomes**

**Coaching and implementation.** It is surprising that neither time nor intensity of coaching was positively associated with implementation; in fact there was a slight negative association between coaching time and W2 FoI scores. However, there are several possible explanations for this finding. First, FoI scores did not differ significantly from W1 to W2, echoing the findings of some other adolescent reading interventions that report no change in teachers’ FoI over the course of a school year (Simmons et al., 2014; Vaughn et al., 2011). This lack of change helps to explain the negative association between FoI and coaching time; as described above, coaches seemed to dedicate more time and energy to teachers whom they perceived as struggling with implementation. Additionally, since all teachers were at least in their second year of implementation and W1 observations didn’t begin until several months into the school year, improvements between W1 and W2 might be expected to be fairly minimal. The purpose of coaching is to facilitate global improvements in instruction, rather than simply increase adherence to a program. Thus, the observation protocol may be too crude a measure to capture the subtle improvements that a teacher might make midway through his or her second, third or fourth year of implementing an intervention. Therefore, FoI, as measured by the observation protocol, may not be the best measure of coaches’ effectiveness in improving teachers’ instructional practices.
Alternatively, the lack of relationship between coaching and W2 FoI may offer implications for effective coach training and practice. STARI coaches were hired at different times, as the number of participating teachers grew over the five-year research period. Although all three coaches were experienced, and they collaborated frequently, their training was largely informal. Biancarosa, Bryk and Dexter (2010) suggest that one reason that their research, in contrast with a number of other studies, found a strong relationship between coaching and student achievement may be because of the extensive training their coaches received – a full year of professional development before beginning to coach. Likewise, Camburn, Kimball and Lowenhaupt (2008) found that coaches were more likely to engage in high-intensity coaching practices if they had received high-quality training and their job expectations were clear. On the other hand, in Kraft and Blazar’s (2016) study, coaches received increased training between the first and second cohort of teachers served, including a mandate to prioritize classroom management in their coaching sessions, and the effects of coaching were poorer in cohort two than cohort one. This suggests that coaches may benefit from increased training and clearer expectations, but only if these directives are aligned with best practices. More research is needed to link coaching practices with student outcomes in order to identify the most high-impact ways for coaches to spend their time.

**Implementation and student outcomes.** Teacher FoI was associated with the number of workbook pages students completed and the number of words students read during the school year, although FoI did not have a significant relationship with growth in oral reading fluency. Specifically, adherence was associated with both workbook completion and the number of words read, while quality and participant responsiveness
were also associated with total words read. Since adherence was the area in which teachers scored highest at both time points, it is promising that adherence is the aspect of FoI associated with both outcome measures.

While other research (e.g., Cantrell et al., 2013; Cantrell et al., 2010; Feldman et al., 2012; Simmons et al., 2014) has demonstrated inconclusive relationships between FoI and more distal measures of student achievement (i.e. standardized tests), the present study demonstrates a relationship between FoI and proximal measures of student learning. Workbook completion and the total number of words read may be considered relatively insignificant as stand-alone indicators of literacy achievement, but other research has demonstrated that each of these proximal measures is associated with longer-term measures. For example, an instrumental variables analysis indicated that STARI workbook completion had a relationship with standardized test outcomes, such that students who completed more of the workbook made greater gains from pretest to posttest, with effect sizes ranging from 0.08 to 0.35 (Kim et al., 2016). This suggests that students’ completion of comprehension workbooks prepared them to achieve on standardized tests, and thus that workbook completion serves as an adequate proxy for literacy gains. Furthermore, significant research has demonstrated that amount of reading is associated with reading achievement (e.g., Allington, 2012; Guthrie & McRae, 2012; Schaffner, Schiefele, & Ulferts, 2013). Therefore, it seems plausible that the associations between teachers’ fidelity of implementation and students’ workbook completion and number of words read suggest a relationship between FoI and student achievement on a larger scale, and that over time, the variation in FoI described in this study will contribute to differences in students’ achievement in comprehension and fluency.
Although the findings of this study point to the importance of more in-depth examinations of coaching and instruction in implementing and evaluating interventions, several limitations are to be acknowledged. First, although the sample size is similar to that of other nuanced analyses of FoI (e.g., Benner et al., 2011; Fogarty et al., 2014; Zvoch, 2012), the small sample limits generalizability. Secondly, inconsistencies in collection of the observation data may limit the interpretability of the FoI measure. The fact that the number of observations per teacher varied widely; that observations began so late in the school year; and that for some teachers, only three months, or approximately a third of the school year, passed between W1 and W2, all limit the likelihood that teachers’ FoI scores would demonstrate relationships with student outcomes. Third, the coaching data was based on self-report, although these real-time weekly logs probably produced more accurate self-report than the retrospective questionnaires relied upon by many studies. Fourth, the reliance on work completion, rather than objective measures of achievement, in assessing student outcomes may constitute a limitation, although this is mitigated by the use of multiple measures (workbook completion and words read), and the demonstrated relationship of these two measures to standardized assessments of reading achievement.

Finally, expert teachers often make adaptations to curricula in order to meet the needs of their students. Therefore, thoughtful deviations from the curriculum as written may actually benefit participants. Chapter 3 of this dissertation presents teachers’ perspectives on the intervention curriculum, including their rationales for making adaptations to the program rather than implementing with fidelity. Teacher insights about
reasons for variation in implementation might help to illuminate the complex relationship between FoI and student outcomes.

**Conclusion**

The present study emphasizes the importance of more in-depth examinations of fidelity of implementation than are typical in literacy intervention research. Overall high average fidelity scores mask significant teacher-level variation – even in the later stages of intervention implementation - which may in turn lead to variation in student outcomes. Furthermore, the present study suggests that coaches’ choices about how, and how much, to interact with teachers seem to depend more on teachers than on coaches. The study points to several factors which may influence teachers’ implementation of an intervention and coaches’ interactions with teachers, and thus may suggest ways in which future interventions may improve in their support of both teachers and coaches.
Table 1

*Descriptive statistics on fidelity of implementation*

<table>
<thead>
<tr>
<th>n</th>
<th>Adherence</th>
<th>Fluency</th>
<th>Comprehension</th>
<th>Quality</th>
<th>Participant</th>
<th>Disruption</th>
<th>Fluency</th>
<th>Comprehension</th>
<th>FoI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Adherence</td>
<td>Adherence</td>
<td></td>
<td>Responsiveness</td>
<td>Dosage</td>
<td>Dosage</td>
<td>Dosage</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>W1</td>
<td>17</td>
<td>14.65</td>
<td>2.47</td>
<td>6.47</td>
<td>0.62</td>
<td>7.44</td>
<td>2.19</td>
<td>6.18</td>
<td>2.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W2</td>
<td>17</td>
<td>14.41</td>
<td>2.09</td>
<td>6.71</td>
<td>0.47</td>
<td>7.29</td>
<td>2.04</td>
<td>7.00</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Results of regression analyses comparing FoI and coaching

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>W1 FoI</td>
<td>Coaching time</td>
<td>W2 FoI</td>
<td>W2 FoI</td>
<td></td>
</tr>
<tr>
<td>Highest degree earned</td>
<td>-8.06 (5.65)</td>
<td>-2.88 (5.10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years experience</td>
<td>2.16 (6.28)</td>
<td>0.11 (0.34)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years teaching STARI</td>
<td>2.16 (6.28)</td>
<td>-3.13 (5.48)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading certificate</td>
<td>14.73* (5.08)</td>
<td>-3.29 (4.43)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1 Fluency Adherence</td>
<td></td>
<td>-3.49** (1.00)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1 Comprehension Adherence</td>
<td></td>
<td>0.90 (1.21)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1 Quality</td>
<td></td>
<td>-1.10 (0.87)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1 PR</td>
<td></td>
<td>1.03 (1.15)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1 FoI</td>
<td></td>
<td>0.24 (0.21)</td>
<td>0.42  (0.22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coaching Time</td>
<td></td>
<td>-0.01* (0.01)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coaching Intensity</td>
<td></td>
<td></td>
<td>-7.56 (7.63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>64.83 (12.83)</td>
<td>32.44 (11.53)</td>
<td>67.23 (10.96)</td>
<td>53.55 (15.05)</td>
<td>44.69 (19.44)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.556</td>
<td>0.122</td>
<td>0.553</td>
<td>0.395</td>
<td>0.252</td>
</tr>
<tr>
<td>N</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
</tbody>
</table>

Note. *p<.05, **p<.01
### Table 3

*Descriptive statistics on coaching*

<table>
<thead>
<tr>
<th>Type of Contact</th>
<th>Quick checkin</th>
<th>Planning/debriefing meeting</th>
<th>lesson observation</th>
<th>modeling/demonstration</th>
<th>coteaching</th>
<th>PLC or network meeting</th>
<th>email or phone exchange</th>
<th>no/minimal contact</th>
<th>other activity</th>
<th>Total level 1</th>
<th>Total level 2</th>
<th>Total level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Intensity</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Coach 1</td>
<td>19 (9%)</td>
<td>25 (12%)</td>
<td>29 (14%)</td>
<td>0</td>
<td>8 (4%)</td>
<td>32 (16%)</td>
<td>21 (10%)</td>
<td>60 (29%)</td>
<td>11 (5%)</td>
<td>36 (34%)</td>
<td>40 (38%)</td>
<td>31 (29%)</td>
</tr>
<tr>
<td>Coach 2</td>
<td>23 (8%)</td>
<td>56 (19%)</td>
<td>32 (11%)</td>
<td>3 (1%)</td>
<td>66 (22%)</td>
<td>21 (7%)</td>
<td>12 (4%)</td>
<td>68 (23%)</td>
<td>14 (5%)</td>
<td>25 (19%)</td>
<td>23 (17%)</td>
<td>84 (64%)</td>
</tr>
<tr>
<td>Coach 3</td>
<td>52 (11%)</td>
<td>64 (13%)</td>
<td>70 (15%)</td>
<td>4 (1%)</td>
<td>54 (11%)</td>
<td>25 (5%)</td>
<td>134 (28%)</td>
<td>63 (13%)</td>
<td>12 (3%)</td>
<td>68 (40%)</td>
<td>26 (15%)</td>
<td>74 (44%)</td>
</tr>
<tr>
<td>Total</td>
<td>94 (10%)</td>
<td>145 (15%)</td>
<td>131 (13%)</td>
<td>7 (1%)</td>
<td>128 (13%)</td>
<td>78 (8%)</td>
<td>167 (17%)</td>
<td>191 (20%)</td>
<td>37 (4%)</td>
<td>129 (31%)</td>
<td>89 (22%)</td>
<td>189 (46%)</td>
</tr>
</tbody>
</table>

**Focus of Contact - Process**

<table>
<thead>
<tr>
<th></th>
<th>lesson planning</th>
<th>routines</th>
<th>adaptation</th>
<th>management</th>
<th>assessments</th>
<th>materials</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coach 1</td>
<td>32 (19%)</td>
<td>40 (24%)</td>
<td>32 (19%)</td>
<td>5 (3%)</td>
<td>24 (14%)</td>
<td>2 (1%)</td>
<td>32 (19%)</td>
</tr>
<tr>
<td>Coach 2</td>
<td>26 (15%)</td>
<td>27 (15%)</td>
<td>23 (13%)</td>
<td>22 (13%)</td>
<td>10 (6%)</td>
<td>6 (3%)</td>
<td>62 (35%)</td>
</tr>
<tr>
<td>Coach 3</td>
<td>28 (13%)</td>
<td>29 (13%)</td>
<td>58 (26%)</td>
<td>18 (8%)</td>
<td>10 (5%)</td>
<td>25 (11%)</td>
<td>52 (24%)</td>
</tr>
<tr>
<td>Total</td>
<td>86 (15%)</td>
<td>96 (17%)</td>
<td>113 (20%)</td>
<td>45 (8%)</td>
<td>44 (8%)</td>
<td>33 (6%)</td>
<td>146 (26%)</td>
</tr>
</tbody>
</table>

**Focus of Contact - Content**

<table>
<thead>
<tr>
<th></th>
<th>Fluency</th>
<th>Word Study</th>
<th>Comprehension</th>
<th>Guided Reading</th>
<th>Partner Work</th>
<th>Debate</th>
<th>Non-content related</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coach 1</td>
<td>24 (18%)</td>
<td>3 (2%)</td>
<td>6 (4%)</td>
<td>38 (28%)</td>
<td>41 (30%)</td>
<td>6 (4%)</td>
<td>10 (7%)</td>
<td>9 (7%)</td>
</tr>
<tr>
<td>Coach 2</td>
<td>36 (19%)</td>
<td>12 (6%)</td>
<td>4 (2%)</td>
<td>31 (17%)</td>
<td>35 (19%)</td>
<td>13 (7%)</td>
<td>5 (3%)</td>
<td>50 (27%)</td>
</tr>
<tr>
<td>Coach 3</td>
<td>29 (12%)</td>
<td>7 (3%)</td>
<td>24 (10%)</td>
<td>59 (24%)</td>
<td>45 (19%)</td>
<td>18 (7%)</td>
<td>31 (13%)</td>
<td>31 (13%)</td>
</tr>
<tr>
<td>Total</td>
<td>89 (16%)</td>
<td>22 (4%)</td>
<td>34 (6%)</td>
<td>128 (23%)</td>
<td>121 (21%)</td>
<td>37 (7%)</td>
<td>46 (8%)</td>
<td>90 (16%)</td>
</tr>
<tr>
<td>Fixed Effect</td>
<td>Workbook completion</td>
<td>Total Words Read</td>
<td>Spring WPM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------------</td>
<td>------------------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>β</td>
<td>SE</td>
<td>β</td>
<td>SE</td>
<td>β</td>
<td>SE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>36.42 (61.01)</td>
<td>192.85 (1980.67)</td>
<td>94.12</td>
<td>(22.97)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FoI 2.49* (1.28)</td>
<td>120.67** (41.21)</td>
<td>0.36 (0.42)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adherence 5.69* (2.38)</td>
<td>254.60*** (77.14)</td>
<td>0.74 (0.82)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality 4.63 (0.23)</td>
<td>280.35*** (126.21)</td>
<td>1.10 (1.17)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant responsiveness</td>
<td>9.61 (0.13)</td>
<td>436.06* (216.12)</td>
<td>0.43 (2.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluency dosage 0.03 (1.62)</td>
<td>-41.58 (56.79)</td>
<td>0.16 (0.48)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension dosage -0.40</td>
<td>-3.17 (18.31)</td>
<td>0.21 (0.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall WPM 0.43*** 0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random Effect</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher level 46.65 (8.39)</td>
<td>1477.19 (271.66)</td>
</tr>
<tr>
<td>Student level 38.16 (1.64)</td>
<td>1667.67 (71.71)</td>
</tr>
<tr>
<td>ICC 0.599</td>
<td>0.440 0.121</td>
</tr>
</tbody>
</table>

Note: *p<.05, **p<.01, ***p<.001
Figure 1. Teachers, nested by school and coach
Partner work
Ninth Ward (pp. 174 – 179)

Read to the on page 175.
What does TaShon offer to do? How does it make Lanesha feel, and why? Talk to your partner.
TaShon offers to __________________________________. This makes Lanesha feel ______
Read to the end of the chapter on page 179.

Find these sentences on page 177:
“...”
Find these sentences on page 177:

I put my hand to her chest, listening for breathing. Feeling for her lungs to rise. It’s only then that I cry.”

Read the whole paragraph where you found these words, either silently or out loud to your partner.
□ I read it silently. □ I read it out loud.
What very important thing happened in this paragraph on page 177?
On page 177, Mama Ya-Ya
How do you know? What information in the text helped you figure this out?
The book said

What information in your head helped you figure this out?
I know that

At the end of the chapter, Lanesha counts steps and figures out when the floodwater will fill the attic of her house. What will she do to survive? Write a prediction:

Day 25

Unit 1.4—Hurricane Katrina 1 p. 81

Figure 2. Sample page from STARI student workbook
Chapter 3.

Teacher Adaptations to an Adolescent Reading Intervention

The Institute of Education Sciences currently funds 22 adolescent literacy intervention studies in the US, collectively affecting tens of thousands of students at a cost of over a quarter of a billion dollars (Institute of Education Sciences, 2016). Most IES-funded literacy interventions involve implementing researcher-designed literacy curricula in partner schools, training teachers in specific instructional strategies, typically over the course of three to five years. After this time, researchers leave the partnerships, and implementation supports are withdrawn. Presumably, the goal in designing and funding these interventions is not only to impact the limited number of students who actually participate in the study during a closed time frame, but to better educate larger populations of students, long after the research is complete. However, to my knowledge, no research has been conducted into the nature of instruction during sustained implementation of an adolescent reading intervention, after the period of the randomized controlled trial during which fidelity of implementation is expected.

The present study attempts to address this gap by taking a detailed look at the instructional practices of four teachers implementing their version of a literacy intervention after the formal conclusion of the intervention study. After the randomized controlled trial of STARI concluded, one of the nine participating schools chose to formally sustain implementation of the intervention at the school level. The four STARI teachers in this school, all of whom were experienced STARI teachers in at least their second year of implementation, were observed and interviewed with a focus on the
adaptations they made to the curriculum. The study is framed around the research questions:

What adaptations do teachers make to the formal curriculum? Do they meet the criteria for productive adaptations? What orientations toward the intervention do these adaptations reflect?

**Conceptual Background**

Contemporary scholars distinguish among the *formal curriculum*, composed of the goals and activities which are written in textbooks or unit plans, the *intended curriculum*, which refers to teachers’ plans and aims, and the *enacted curriculum*, which describes the teaching and learning that actually take place in the classroom (Remillard, 2005). In the present study, the *formal curriculum* refers to a yearlong sequence of literacy unit plans, each composed of a series of daily lesson plans, focused on a discrete and thematically linked set of texts and skills. The formal curriculum is never identical to the enacted curriculum; instead, teachers and formal curricula interact to produce the enacted curriculum (Ball & Cohen, 1996). This interaction takes place through teachers’ adaptations of the formal curriculum, in other words, adding, omitting, modifying or substituting instructional activities (Blakely et al., 1987; Drake & Sherin, 2006; Forbes, 2011). Adaptation is inevitable, as all teachers make changes between formal and enacted curricula (Burkhauser & Lesaux, 2015; Datnow & Castellano, 2000; Leko et al., 2015). For example, a study of the elementary reading program Success for All (SFA) noted that not only did all teachers adapt the curriculum, but teachers’ adaptations were not associated with their level of support for the program; those who merely “accepted” SFA made no more changes during implementation than those who supported it (Datnow & Castellano, 2000).
In addition to being inevitable, adaptations may be desirable; Durlak and DuPre’s (2010) review of the research found that some adaptations improved program outcomes. Adaptations may be especially critical in considering sustained implementation of an intervention curriculum. Datnow and Castellano (2000) conclude their study of SFA with the following questions:

how can schools achieve ownership among teachers for an externally developed reform?...Does creating ownership require some level of local adaptation or development? If some level of adaptation is inevitable and even necessary to engage teacher support and program sustainability, how much adaptation is really too much? At what point do individual teachers' adaptations compromise the quality of implementation school-wide?...These questions are critical for the institutionalization of SFA and other externally developed reform models and thus merit further study (p. 795).

Fifteen years later, these questions still have not been addressed in depth in the field of literacy. The majority of research on curriculum adaptation has taken place in math (e.g., Bergqvist & Bergqvist, 2016; Choppin, 2011; Drake & Sherin, 2006; Hill & Charalambous, 2012; Remillard, 2000, 2005; Simon & Tzur, 1999) and science (e.g., Davis et al., 2014; Debarger et al., 2016; Forbes, 2011; Penuel et al., 2014). With a few exceptions, existing research into literacy teachers’ adaptations has focused on teachers who are resistant to a mandated curriculum (e.g., MacGillivray, Ardell, Curwen, & Palma, 2004; Maniates, 2016; Palmer & Rangel, 2011; Pease-Alvarez, Samway, & Cifka-Herrera, 2010; Stillman & Anderson, 2015), teachers who are new to the field (e.g., Grossman & Thompson, 2008; Pardo, 2006; Valencia, Place, Martin, & Grossman, 2006) or both (e.g., Achinstein & Ogawa, 2006; Smagorinsky, Lakly, & Johnson, 2002).

The few exceptions include Leko, Roberts and Pek’s (2015) analysis of adaptations made by five teachers implementing a computer-based reading intervention for middle schoolers (System 44), Burkhauser and Lesaux’s (2015) analysis of six middle
school teachers’ adaptations to a vocabulary curriculum (ALIAS), and Datnow and Castellano’s (2000) analysis of implementation of SFA in two schools. All three of these studies conducted classroom observations and interviews with teachers. Collectively, these studies found that all teachers made adaptations in response to students’ perceived cognitive, affective and behavioral needs, as well as in response to teachers’ own strengths, experiences and preferences. For example, three of the System 44 teachers were trained as secondary English teachers, with little training in basic reading instruction. Thus, they supplemented the prescribed phonics program with instruction in vocabulary and comprehension, which they believed to be their areas of strength as educators (Leko et al., 2015). As another example, some SFA teachers acknowledged that they made adaptations to make teaching easier or more enjoyable (Datnow & Castellano, 2000). None of these studies addressed sustained implementation of an intervention, after research-related supports had been withdrawn. In fact, in two of these studies (Burkhauser & Lesaux, 2015; Leko et al., 2015), teacher participants were implementing the curriculum for the first time. Thus, the present study makes an important contribution to the literature by examining sustained implementation by experienced teachers.

**Sustained Implementation**

To my knowledge, the only study conducted into sustained implementation of a literacy intervention, after the conclusion of the formal trial, investigated teachers’ sustained use of Kindergarten Peer Assisted Learning Strategies (KPALS) (Kearns et al., 2010). This study surveyed teachers one year after the conclusion of the formal study, and based their definition of sustained use on teachers’ response to a single item inquiring
whether they were still implementing KPALS. No classroom observations were conducted to analyze this implementation in greater depth.

Yet sustained implementation is critically important in order to build teachers’ expertise. Significant research suggests that student achievement improves as teachers gain experience in implementing a curriculum (Biancarosa et al., 2010; Lovett et al., 2008), and that teachers are more likely to make adaptations in later iterations of curriculum implementation (Debarger et al., 2013; Drake & Sherin, 2009; Grossman & Thompson, 2008). Two of the teachers in the present study were in their second year of implementing STARI, and two were in their fifth year. Thus, the adaptations described in the present study may be considered “learned adaptations,… knowledge-based adaptations designed with respect to what teachers have learned from prior enactments” of curriculum (Choppin, 2011, p. 335). By focusing on experienced teachers implementing learned adaptations to a literacy curriculum, the present study addresses an important, yet under-researched, area of curriculum enactment.

**Adaptation and Orientation toward Curriculum**

In order to understand teachers’ enactments of curriculum, it is essential to understand teachers’ preexisting orientations toward the curriculum (Spillane, Reiser, & Reimer, 2002). Remillard and Bryans (2004) define orientation toward curriculum as “a set of perspectives and dispositions about [subject matter], teaching, learning and curriculum that together influence how a teacher engages and interacts with…curriculum” (p. 364). After observing eight teachers four or more times each over the course of two years, Remillard and Bryans created a name for each teacher’s
orientation toward the mathematics curriculum under investigation, i.e., adherent and trusting, skeptical or quietly resistant.

The present study, however, views teachers’ orientation toward curriculum not as static, but as a process of sense-making (Coburn, 2004; Spillane et al., 2002). As Remillard and Bryans (2004) argue, preexisting orientations may act as a “lens” through which a teacher views a new curriculum, influencing which aspects of the curriculum the teacher attends to, and how those aspects “are processed, encoded, organized, and subsequently interpreted” (Spillane et al., 2002, p. 394). Yet, encounters with new teaching methods may also change teachers’ orientations toward curriculum. The present study analyzes teachers’ orientations toward curriculum after two to five years of implementation, considering the ways in which their preexisting orientations have shaped their enactments, but also considering the ways in which their orientations toward literacy teaching and learning may have changed as a result of interacting with the curriculum.

In her study of teachers’ sense-making processes in response to policy messages, Coburn (2004) proposed that teachers may respond in five possible ways: the first is rejection, wherein the teacher does not implement any part of the reform. Secondly, in symbolic response or decoupling, the teacher makes alterations to appear as though he or she is implementing a new policy without making any substantive instructional changes. In parallel structures, a teacher implements new practices alongside existing practices; in assimilation, the teacher adopts aspects of an initiative but transforms them to fit his or her underlying beliefs or practices; and finally, in accommodation, a teacher changes his or her beliefs and practices in accordance with the new policy. The present study examines teachers’ orientations toward the curriculum in light of this framework,
considering which of these five responses teachers may be demonstrating. When considering orientation, this study takes an actor-oriented perspective toward curriculum implementation, focusing on teachers’ decision-making and rationales, rather than on fidelity to the formal curriculum (Penuel et al., 2014).

**Productivity of Adaptations**

Despite this actor-oriented focus on curriculum adaptation as a process of sense-making, “it would be inaccurate and irresponsible to conclude that all interpretations of a written curriculum are equally valid” (Remillard, 2005, p. 239). Therefore, the third research question of the present study focuses on a limited number of curriculum adaptations – three each from two teachers – and presents an in-depth analysis of the productivity of each adaptation to better understand whether the adaptations made by these experienced teachers constitute valid interpretations of the STARI curriculum. The adaptations’ productivity is evaluated based on three criteria that Debarger and colleagues developed when reviewing the literature on curriculum adaptations in math and science (Debarger et al., 2013). First, according to Debarger and colleagues, a productive adaptation must be responsive to multiple stakeholders, meaning that it should be faithful to the intentions of the curriculum designers, while also responding to the needs of students. Secondly, a productive adaptation must incorporate responsive discourse practices. This means that for an adaptation to be considered productive, teachers must move away from the traditional Initiate-Respond-Evaluate in which classroom talk consists of a series of brief dialogues between the teacher and selected individual students (Cazden & Beck, 2003), toward a structure that “emphasize[s] the contributions of all students, encourage[s] students to develop their ideas and listen to
others’ ideas, and support[s] students as they build logical connections and draw reasonable conclusions” (Debarger et al., 2013, p. 302). Teachers may encourage these responsive discourse practices through productive talk moves, such as pressing for reasoning when students make assertions (Michaels & O'Connor, 2015). Finally, Debarger and colleagues require that a productive adaptation maintains or enhances task complexity. For example, if students struggle with a task, a teacher might respond in a way that lowers the cognitive demand (i.e. giving students the correct answer) or that maintains it, for example by asking students to jot down their reasoning before engaging in discussion, providing more thinking time as a scaffold. According to Debarger and colleagues, only the second of these adaptations would be considered productive.

While Debarger et al. (2013) developed this framework for math and science, their criteria are appropriate for literacy curricula, specifically STARI, as well. The first criterion, responsiveness to multiple stakeholders, is arguably applicable to all instruction, as teachers must constantly balance the expectations of curricula and standards with the needs of the students in their classroom. The second criterion, responsive discourse practices, is appropriate to the present study as significant research has demonstrated the importance of open discussion in literacy development (Applebee, Langer, Nystrand, & Gamoran, 2003; Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009; Nystrand, 1997), and consequently student talk is a key component of STARI. Finally, STARI’s focus on selecting texts and activities that balance accessible readability levels with cognitively challenging content makes the third criterion, maintaining or enhancing task complexity, also suitable for this analysis.
In sum, the present study makes a number of important contributions to the research base: first, it presents a nuanced description of teacher practice that is rare in literacy intervention research. Secondly, this study offers a look at sustained implementation of an intervention curriculum, which is almost unheard of in any content area. Furthermore, a close analysis of the productivity of learned adaptations will help us understand “how much adaptation is really too much” (Datnow & Castellano, 2000, p. 795). This embedded single case-study design constitutes a revelatory case, meaning one in which, due to the lack of previous research on this topic, the descriptive information alone can generate important new knowledge of a phenomenon (Yin, 2014). Therefore, the present study conducts a detailed analysis of a small number of teachers enacting a small number of lessons, in order to deeply understand the ways in which teachers adapt curriculum and the implications of these adaptations for research and practice.

**Methods**

**Participants**

Data were collected in sixth and seventh grade STARI classes in one school with a diverse, low-income student population (see Table 1 for school-level student demographics). All STARI classes were small, with between four and ten students, and they met three or four days per week for a full 50-minute class period. All four STARI teachers at the school were observed and interviewed. Ruth is a Caucasian female, a certified reading specialist with 23 years of teaching experience. She was one of the original STARI teachers who participated in piloting the curriculum, and at the time of the study was in her fifth year of teaching STARI. In addition to STARI, she taught Wilson Reading, a scripted phonics program intended as an intensive intervention for
students with disabilities (Wilson Language Training Corporation, 2017). Christine, a Caucasian female with over 25 years of teaching experience, was also in her fifth year of teaching STARI. She held a general elementary teaching certificate. In addition to STARI, she taught Fountas and Pinnell’s Leveled Literacy Intervention (LLI), a small-group literacy intervention for struggling readers (Fountas & Pinnell, 2015). Wendy is a Caucasian female with 31 years of teaching experience, who was in her second year of teaching STARI. She held a general elementary teaching certificate, and prior to STARI had taught mostly fifth and sixth grade, all subjects. She also taught LLI in addition to STARI. Mary is a Caucasian female with 37 years of teaching experience, who was in her second year of implementing STARI. She was formerly a gifted and talented teacher, and held a certificate as a consulting teacher of reading. She also taught LLI in addition to STARI.

All teachers were trained in STARI implementation at a multi-day summer institute before they began implementation, and received one-on-one coaching in their implementation of STARI in each year preceding the 2015-16 school year. They also received detailed daily lesson plans for each unit (see Figure 1 for a sample lesson plan).

Measures

Observations. Each teacher was observed three times over the course of the school year, between December and June. Observations were scheduled at the teacher’s convenience, and thus were announced ahead of time. Lessons were video and/or audio recorded and transcribed for analysis.

Interviews. Immediately following each lesson, each teacher engaged in a brief informal interview about her goals for the day’s lesson, reflections on how the lesson
went, and reasons for any adaptations. After all three lessons had been observed and the enacted lessons had been transcribed and compared with the lesson plans from the formal curriculum, each teacher participated in a fourth, longer interview. In this interview, the interviewer described specific adaptations that had occurred and provided memory aids, including transcript excerpts from the enacted lessons, excerpts from lesson plans, and artifacts such as workbook pages. Teachers were asked to describe their rationale for each adaptation.

This study uses a modified version of Simon and Tzur’s (1999) lesson set methodology, wherein “the unit of analysis is a set, which consists of at least two [lessons]…and interviews before, between, and after the lessons” (p. 258). Therefore, each teacher’s three lessons and four interviews were treated as a set for the purposes of analysis. This method avoided placing excessive emphasis on the events of any one lesson, but instead sought patterns across the set. Although the relatively small number of observations per teacher may constitute a limitation of the study, the number of observations is consistent with that of other curriculum implementation research (e.g., Chval, Chavez, Reys, & Tarr, 2009; Debarger et al., 2016; Leko et al., 2015), and like these researchers, I found that data appeared consistent across observations. In addition, in the final interview, I asked teachers whether the three classes I had observed were typical examples of their instruction, and all four teachers said that they were.

Analyses

Following Simon and Tzur (1999), each lesson set was analyzed to create an account of teaching practice for each teacher. Unlike traditional case studies, accounts of practice are “based on an explicit commitment to understanding the coherence in the
teacher’s practice… and the underlying perspective of the teacher that it potentially reveals” (p. 262). Practice is defined as “not only everything teachers do that contributes to their teaching (planning, assessing, interacting with students) but also… teachers’ intuitions, skills, values, and feelings about what they do” (p. 253-254). An account of practice is created through an analysis of each lesson set with the goal of creating a coherent, theoretically driven representation of a teacher’s work, attending to the specific aspects of practice most relevant to the research questions. Thus, an account of practice “explain[s] the teacher’s perspective from the researchers’ perspective,” (Simon & Tzur, 1999, p. 254; emphasis in original). In other words, although an account of practice analyzes teachers’ perspectives, the researchers’ theoretical focus may lead them to attend to aspects of practice different from those that the teacher herself finds most relevant. Thus, the goal of an account of practice is not to provide an ethnographic description of a teacher’s work, but to suggest a coherent representation of her practice that is “appropriate (given the data),” and relevant to the research questions (Simon & Tzur, 1999, p. 255).

In the present study, analysis of lesson sets began with a side-by-side comparison of the enacted curriculum with the formal curriculum, comparing the transcript of each lesson with the corresponding lesson plan from the formal curriculum. I used Dedoose qualitative coding software (2016) to identify each addition, modification and omission that the teacher made, as well as any activities that she implemented in an unadapted form. I identified all adaptations, even those that seemed obvious, natural or insignificant – for example, keeping the whole class together instead of dividing into groups when the number of students was small, or asking follow-up questions to help clarify students’
thinking or guide them toward an answer. I focused on even these minor adaptations for two reasons: first, “the qualitative research approach demands that the world be examined with the assumption that nothing is trivial [and]…nothing is taken for granted” (Bogdan & Biklen, 1998, p. 6). Even seemingly “trivial” adaptations may provide critical information about a teacher’s orientation toward the curriculum. Secondly, each teacher made different “taken for granted” adaptations (e.g., asking different follow-up questions within the same lesson), highlighting the ways in which each teacher’s enacted curriculum differed from the others’ as well as from the formal curriculum.

After all adaptations had been identified, I used open coding to name or describe each adaptation, creating codes which were grounded in the data and allowing new themes to emerge (Emerson, Fretz, & Shaw, 2011). Next, I used these codes to create focused codes, or more conceptual codes which described larger chunks of data and then recoded the data using these focused codes (see Table 2 for selected codebook) (Charmaz, 2010). For example, the most common adaptation was for teachers to add additional questions that were not written in the formal curriculum. I coded these questions in several ways. For instance, in Christine’s first recorded lesson, on The Skin I’m In (Flake, 2000), the lesson plan poses the question: “What about Desda’s opinion? Can someone read what she thinks, on the top of page 73?” Christine asked her students: “Why would Desda even say that, ‘I wish someone would kill themselves over me?’” I coded this as unadapted, since it is a paraphrase of a question from the formal curriculum. However, when students didn’t get the answer to the question right away, Christine asked further questions, such as “Is she popular, Desda?...What was she good at?” I coded these as addition: follow-up question, since these questions were prompted
by and related to a question from the lesson plan. In contrast, earlier in the same lesson, Christine asked, “Does that give you a clue about John-John?,” a question which was unrelated to any part of the formal curriculum. Therefore, I coded that question as addition: new question. Four of the twelve lesson transcripts were collaboratively coded (Smagorinsky, 2008). In other words, after developing a preliminary codebook, I collaborated with a colleague to code four randomly selected lessons, discussing each data segment until we agreed on how it should be coded, and revising the codebook as necessary. After collaboratively coding these four lessons, I independently coded the remaining eight.

To calculate the frequency of various types of adaptations, I divided each lesson transcript into five-minute segments, and graphed which adaptations were present during each segment. This segmenting helped to appropriately “weight” each adaptation; for example, an added activity lasting twenty minutes and a single follow-up question could not each be fairly counted as one adaptation. Thus, the frequency of each type of adaptation was determined by summing the number of five-minute segments in which that adaptation occurred. Next, I conducted a cross-case analysis to seek patterns across teachers. I created a matrix listing the frequency of each adaptation (e.g., addition: new question) by teacher to identify (1) which adaptations were most vs. least common overall, and (2) which adaptations were made relatively equally often by all teachers, as opposed to those made disproportionately by some teachers, but not by others.

After identifying and coding all adaptations in the lesson transcripts, I turned to the interview portion of the lesson sets. I began by identifying all excerpts of text in which teachers discussed adaptations. Following a similar process to the above, I then
used open coding to try to name what the teacher considered when making each adaptation. For example, when teachers talked about increasing the amount of student discussion by adding questions, as described above, they offered both *engaging students* and *building comprehension* as reasons for the importance of discussion – rationales that reflect different processes for making sense of the formal curriculum.

To address my second research question about the productivity of adaptations, I chose to focus on two teachers (Ruth and Mary) because their adaptations entailed more substantive alterations to the formal curriculum than those of the other two teachers. I then selected three focal adaptations for each teacher that represented substantial departures from the formal curriculum. In selecting these focal adaptations, I wanted to choose a small enough number to facilitate in-depth qualitative analysis, while avoiding excessive focus on exceptional or idiosyncratic teaching. My decision to focus on two teachers, three adaptations each, allowed me to avoid placing excessive emphasis on the events of any one classroom, or any single instructional interaction, while maintaining a manageable corpus of data for this analysis. After selecting the six adaptations, I reread each teacher’s four interviews and excerpted portions where teachers discussed focal adaptations, in order to understand the teacher’s rationale for making the adaptation. Then I engaged in theoretical coding (Willig, 2013; Yin, 2014), comparing each adaptation to each of Debarger et al.’s (2013) criteria for productivity.

Finally, to address my third research question, I worked to develop a coherent understanding of each teacher’s orientation toward the STARI curriculum. Some previous research has named or identified teachers’ orientations toward the particular curriculum under investigation; for example, Datnow and Castellano (2000) categorized
the 36 teachers in their sample along a continuum from *strong supporters* to *vehemently against* the SFA reading program. Remillard and Bryans (2004) created phrases to describe their eight teachers’ perspectives on the math curriculum they studied, including *adherent and trusting*, *quietly resistant*, and *skeptical*. However, I chose to break this process into two steps: in the first phase of my analysis, I used data from the lesson sets to name, from the researcher’s perspective, what each teacher was oriented toward, in other words, what seemed to each teacher to be the most salient feature(s) of her literacy instruction. For example, a teacher might be oriented toward the content of the assigned texts, or toward fidelity to the STARI program, as the most important aspect of her instruction. After naming each teacher’s orientation toward the curriculum, I engaged in theoretical coding (Willig, 2013; Yin, 2014), comparing this orientation with Coburn’s (2004) framework, described above, to determine what response to the intervention was signaled by her adaptations. This two-step process allowed me to consider each teacher’s orientation toward literacy teaching in general, outside of this specific program, and then to analyze how this orientation shaped and was shaped by the teacher’s implementation of STARI, specifically.

**Findings**

**RQ1: Description of Adaptations**

Of the 118 total 5-minute segments, only 16 (14%) unfolded exactly according to the lesson plan. In other words, of approximately ten hours of recorded lessons, only 80 minutes did not contain some level of adaptation. Categories of adaptation are presented descriptively in Figure 2, which illustrates that additions were the most common type of adaptation, occurring in 110 of the 118 total recorded segments. Modifications,
omissions, and implementing STARI in an unadapted form were all approximately equally frequent, occurring during 69, 58, and 62 recorded segments, respectively. This figure also illustrates that the four teachers were not equally likely to make each type of adaptation. For example, Christine made the largest proportion of the additions, Mary made more modifications and omissions than the other teachers, and Ruth was most likely to implement the curriculum in an unadapted way.

Going beyond these high-level categories, the most common specific adaptation was for teachers to change the questions that were asked, both by omitting questions suggested in the formal lesson plans and by adding original questions. The twelve lesson plans included a total of 41 questions for teachers to pose to students about the texts, and teachers asked 29 of these questions. They were asked entirely by Wendy, Ruth and Christine; Mary never asked a question from the lesson plan. Teachers asked additional questions that were not in the lesson plan during 32 of the 118 five-minute segments (27%). While the four teachers were equally likely to add follow-up questions (each added follow-up questions during five lesson segments), Christine was responsible for more than half of the new questions (she added new questions during 13 lesson segments, compared to Ruth’s six, Wendy’s two and Mary’s one).

The next most common adaptation was for teachers to change the mode of reading in which students engaged. Every STARI lesson plan called for both guided and partner reading. In the three recorded lessons, Ruth always followed the guidance of the lesson plan for mode of reading, but the other teachers never did. Christine always had her whole group sit in a circle and take turns reading the text aloud round-robin style. Wendy reported that on alternating days, she assigned her students to read independently,
or gave them the choice of reading independently or with a partner; I observed both of these practices. Mary’s students did not engage in a first reading of any text during the three recorded lessons.

Additional adaptations included adding teacher-created activities to the lesson plan or modifying the use of materials such that the purpose of the activity changed. For example, Ruth added a prop (a red and blue rope) to provide a concrete illustration of conflict. Mary adapted one activity, in which students were supposed to cut and paste the words of a poem to match the formatting in the text, into a creative writing activity, where students used the words to create their own found poems. Another frequent adaptation was to add a review of the previous lesson’s work. Ruth and Mary each did this once, but Christine and Wendy included this type of review in every recorded lesson. In Wendy’s lessons, in particular, between twelve and twenty minutes was spent reviewing each day – a substantial chunk of a 50-minute period. Finally, an additional common adaptation was to omit workbook pages assigned by the formal curriculum. For example, across Ruth’s three recorded lessons, the formal plans assigned five workbook pages, but Ruth had her students complete only one. Christine also assigned one of five pages from the formal plans, Wendy assigned two of six, and Mary assigned one of eight.

The activity most commonly implemented in unadapted form was fluency practice. Of the twelve recorded lessons, nine lesson plans in the formal curriculum began with fluency practice. In four of these lessons, as well as in one where the plan didn’t include fluency, teachers had their students engage in fluency practice exactly as it was laid out in the formal curriculum. Mary did not include fluency practice in any of her three recorded lessons, but Ruth, Wendy and Christine all included fluency practice one
or two times each (and each omitted it once when it was included in the plan). However, in all five cases in which fluency practice occurred, the teachers added a debrief element, during which students shared and discussed the contents of the passages they had read. In most of these instances, the discussion was fairly brief, but in one of Ruth’s lessons, the fluency debrief lasted nine minutes, and in one of Christine’s, the debrief lasted eleven minutes.

**RQ2: Productivity Analysis**

To address my second research question, on productivity of adaptations, I chose to focus on Ruth and Mary. Although these two teachers did not make the greatest number of adaptations, the adaptations that they made seemed to depart most substantively from the formal curriculum. For each teacher, I selected three adaptations to analyze for productivity according to Debarger and colleagues’ (2013) three criteria: responsiveness to multiple stakeholders (i.e., curriculum designers and students), responsive discourse practices, and maintaining or enhancing task complexity.

All three of Ruth’s focal adaptations took place during two lessons from Unit 1, for which the core text is *The Skin I’m In* (Flake, 2000). In this book, the main character, seventh grader Maleeka, is bullied for being dark-skinned and for wearing homemade clothes sewn by her mother. Maleeka makes a deal with a bully named Char that Maleeka will do Char’s homework in exchange for borrowing Char’s fashionable clothes. The novel depicts the escalation of Char’s bullying (e.g., Char forces Maleeka to get her a new hamburger in the school cafeteria, Char shoves a cigarette into Maleeka’s hand in the school bathroom) until the climax of the story in which Maleeka and Char light a fire to destroy teacher Miss Saunders’ classroom. Ruth was recorded teaching a group of
seven sixth graders, four boys and three girls (although in her second recorded lesson, only five students – four boys and one girl – were present).

**Adaptation 1: Increasing complexity and adding scaffolds.** Ruth’s first adaptation occurred during a lesson that took place in December of 2015. The lesson plan called for students to engage in fluency practice and then read two chapters of the novel and complete two workbook pages. The first page in the workbook asked students to find and copy three quotes from the text describing the character Momma (Maleeka’s mother), and on the next page, students were supposed to draw a picture of Momma.

Ruth adapted this lesson in several ways. She chose to skip fluency practice, to have students read only one chapter and to complete the first workbook page (identifying evidence from the text) but not the second (drawing a picture), and she added an activity focused on character traits.

For this added character trait activity, Ruth provided two tools that were not a part of the STARI curriculum. The first was a list of words students could use to describe a character (see Appendix B), and the second was a blank T-chart for pairing character trait words with textual evidence, both of which she had been given by one of the English teachers at the school. At the beginning of the class period, she introduced the character trait list, saying, “this is something we haven’t done before, but I think some of you might be using that list in Ms. R's class.” The students then read the chapter and completed the workbook page. When she brought the class back together, they briefly discussed the vocabulary term “don’t have all her marbles,” and Ruth segued smoothly into use of the character trait list. A student, Haley, said, “She's a person that likes to get along with kids, and she doesn't want to act like an adult I guess.” Ruth responded,
indicating the list: “Let's see if we can find some character traits here. She doesn't act like an adult. What would you say she acts like? What's another word? Childish. She's childish. She's also what?” Students responded with more words from the list. Many of the words were accurate descriptions of the character, like brilliant, loving and hopeful. When students offered these suggestions, Ruth repeatedly pressed for reasoning (Michaels & O'Connor, 2015), with comments like: “Where is the evidence of this?” and “Tell us why you think she’s hopeful.” When students suggested character traits that were off base, Ruth engaged with their thinking rather than providing negative evaluation, as illustrated in her response to a student’s characterization of Momma as “insistent:”

*Ruth:* Insistent? I don’t know. Is she insistent? Does she insist on anything?
*Unknown Student:* I don't know what insist means.
*Ruth:* It means to make you do something. I insist you do your homework, right? You need to do your homework. I insist. Is she strict?

After a seven-minute discussion using the character trait list, Ruth segued into an activity using the T-chart. She modeled creating a T-chart on the board, with the trait “smart” on the left side, and the textual evidence, “She’s a math whiz and can add numbers faster than anyone I know” on the right. Students copied this onto their T-chart handouts. At that point, class was over, but Ruth told students: “Tomorrow we’ll do a couple together.”

*Ruth’s rationale.* In interviews, Ruth explained that she decided to incorporate these tools because “it kind of fits perfectly into what I'm doing, and it's reinforcing what they're doing in ELA [English language arts] class,” and because “I think that having the list helped them – so they didn't have to do so much thinking, sort of like having the answers at the top of the page before the questions. So I think that's really helpful, especially for kids who don't have good recall or don't have good vocabulary. They don't
have to think about it. It's right there.” Therefore, her decision to incorporate this adaptation seemed to be based on desires to support and reinforce her students’ ELA classes, and to scaffold the activity for her students.

**Responsive to multiple stakeholders.** The stated objective for the lesson plan was for students to “find evidence in a text to support judgments about character and plot,” an objective that Ruth’s activity clearly addressed. Furthermore, STARI’s theory of change posits that student talk is the mechanism through which students will build deep comprehension skills. Thus, engaging students in a discussion of Momma’s character traits, facilitated by Ruth’s productive talk moves, is clearly faithful to the intentions of the curriculum designers. In addition, it seems that Ruth intended this activity to respond to student needs, as she stated a belief that “having the list helped them.”

**Responsive discourse practices.** Within this adaptation, Ruth engaged in responsive discourse with students, repeatedly pressing for reasoning in response to answers that were on- and off-track. She encouraged students to contribute and to develop their ideas, and supported them in drawing reasonable conclusions (Debarger et al., 2013). Thus, this adaptation clearly meets this criterion for productivity.

**Task complexity.** Despite Ruth’s statement that she made this adaptation so students “didn’t have to do so much thinking,” a comparison of this adaptation with the formal curriculum suggests otherwise. The lesson plan required students only to identify and copy three quotes from the text to describe Momma, without any requirement to infer character traits from this textual evidence. Ruth’s adaptation required students to infer character traits (with the list as an aid) and to provide textual evidence to support them. This adaptation clearly enhances the complexity of the task.
Adaptation 2: Increasing complexity without added scaffolds. Ruth’s second adaptation took place during a lesson on *The Skin I’m In* (Flake, 2000) recorded in January 2016. In this lesson, students were supposed to read and discuss two chapters of the text, including the climax of the novel. Instead, Ruth began the lesson by reviewing the concept of conflict, then had students read one chapter and engage in discussion. Ruth then led the students in writing a brief summary of the chapter, then reviewed conflict a second time, and then had students begin their homework. I will consider two of these adaptations: the summary students wrote and Ruth’s teaching of conflict.

After Ruth’s students had read the chapter, and engaged in discussion using questions from the lesson plan, she directed students to write a “ten-word gist,” or brief summary of the chapter. She distributed sticky notes for the students to write on, and modeled writing her own gist on a small white board. She told me that students completed a ten-word gist after every chapter, and students were clearly familiar with the activity; for example, they knew Ruth’s rule that “names don’t count” towards the ten-word maximum. As in the character trait activity, Ruth used questioning to address misunderstandings and to push students to add information. For instance, when student Thomas suggested, “Maleeka got her shoulder blade broken and busted her knee open,” Ruth responded: “Wait a minute. Is that the important part? What happened that was the most important part here?” When Haley suggested, “Maleeka, Char, and twins got caught by janitor,” Ruth responded, “Got caught doing what?... How did they destroy the room? What's the biggest part of it?” Ruth and the students spent approximately two minutes completing this activity.
**Ruth’s rationale.** Ruth stated that the ten-word gist is her preferred summarizing activity because “they have to synthesize what they read.” Furthermore, she added that students would place the sticky notes with gists written on them in the text at the end of the chapter to use as a reference and a “memory thing.” She also said several times that her students “really like doing” the gist. Therefore, she made this adaptation to encourage her students’ ability to synthesize text into a brief summary, to provide a record of what they had read that they could consult in the future, and perhaps to engage her students.

**Responsive to multiple stakeholders.** The objective for this lesson plan was for students to “summarize using the 5 W’s.” Ruth’s adaptation, having students summarize using a 10-word gist, seems to be faithful to the intention of the curriculum designers, as it is simply a different activity aimed at the same skill. Also, Ruth had clearly considered how this adapted activity might benefit her students.

**Responsive discourse practices.** This particular adaptation did not appear to be related to responsive discourse practices. Students worked largely independently (though several spontaneously shared their work aloud). Ruth engaged with their thinking, but largely in the form of veiled corrections (“Wait a minute. Is that the important part? What happened that was the most important part here?”) and requests for additional information (“Got caught doing what?... How did they destroy the room? What's the biggest part of it?”).

**Task complexity.** Ruth argued that a ten-word gist, as compared to a 5 W’s summary, would force students to synthesize the text they had read. While this seems likely, given the level of support students needed to complete the task – essentially, through questioning, Ruth told them what to write – it is possible that they would have
derived more benefit from the less complex task afforded by the 5 W’s summary.

Furthermore, of the 41 conversational utterances that took place during the completion of this task, fifteen (or more than a third) were solely concerned with the number of words a student had written (e.g., “Isn’t that only one word ’cause the names don’t count?” “You should count ‘&’ even though it’s a symbol.”) This adaptation clearly had the potential to increase the complexity of the task; however, this enactment of the adaptation may not have done so effectively.

**Adaptation 3: Adding a tool and creating confusion.** In the same lesson, Ruth added instruction around the concept of conflict, using a long piece of colored yarn – red on one side and blue on the other – which she had draped around her neck like a scarf.

She began class by allowing students to participate in summarizing the chapter they read in the previous lesson, then interjected, “Oh, so there's a conflict. Okay, so Maleeka is conflicted because before this – so you see this? This is the rope. It's like a tug of war… You see it?… This is conflict, okay? It's the tug of war.” Through questioning, she tried to get students to see that characters Caleb and Char are pulling Maleeka in opposite directions. However, as illustrated in the following excerpt from the lesson transcript, the students seemed to struggle to come up with the answers she was looking for, and Ruth ended up answering most of her questions herself:

*Ruth:* What's Caleb's plan?
*Kyle:* He's like trying to convince her not to go.
*Ruth:* What does he want to do though?
*Haley:* Well, Maleeka likes –
*Thomas:* He's trying to change things.
*Ruth:* In the school, right?
*Thomas:* Yeah.
*Ruth:* He wants to – he was cleaning the bathrooms. He wants to clean things up, right? So that's Caleb pulling at her. She's that knot in the middle. Okay?
Thomas then brought up another character, and Ruth tried to prevent the conversation from deviating.

Ruth: Okay, so we have Caleb. I want to get back to Caleb. So Caleb is in the red, right? He's saying –

Roman: Don't go.

Ruth: Well, he's saying – he says – I'll quote it for you, "Char's the kind of friend that will get you locked up or shot up." Ooh.

Roman: Shot up?

Thomas: Oh, wow.

Ruth: And he wants to clean up this McClenton High School, right? What about Char? She's the blue.

Ethan: She wants to like –

Ruth: No, she's planning on doing what to the school or the room?

Kyle: To jack up Miss Saunders and the room.

[The class then briefly discusses Char.]

Ruth: So do you see the conflict that Maleeka's having? Does she do a good thing or does she do the bad thing?

Roman: I feel like she's right in the middle.

Ruth: She's right in the middle. What do you think she's going to do?

[The class then makes predictions before they begin to read the chapter. Conflict is not mentioned again before reading.]

The transcript excerpt above illustrates some challenges that Ruth faced incorporating this additional tool – the colored rope – into the lesson. She did not engage with student thinking when they suggested answers different from the ones she had in mind (e.g., when Kyle said, “He’s trying to convince her not to go,” or when Ethan said, “She wants to like-“ and Ruth cut him off). While there did seem to be some level of student understanding by the end of the discussion (“She’s right in the middle”), it was unclear whether the rope tool contributed to students’ understanding. This activity took approximately two minutes.

After students read the chapter, discussed it, and completed the ten-word gist, Ruth returned to using the rope as a tool to support students’ understanding of conflict. It appeared that she did so on the spur of the moment, based on having more time than expected left in the lesson, as she introduced the activity by saying: “I'm going to give you some choices here, okay? I didn't plan on this, but I'm going to give you some
choices….Are you understanding the narrative arc a little bit better?... Do you want me to go over that again? Let me just go over it.” She began by pointing at a narrative arc chart (see Appendix B) on the whiteboard, and then referred to the rope, asking students to list conflicts that had taken place in the story so far, and tying a knot in the rope to represent each conflict. The class came up with a total of seven conflicts, of which three were suggested by students and four by Ruth. Typically, students were able to come up with the conflict Ruth had in mind through her questioning. For instance, Ruth asked: “Did [Maleeka] ever try to get back at Char?” Students replied, “No.” Ruth prompted, “Secretly? Come on. Think about it.” Thomas then came up with the example that Ruth seemed to have in mind. Throughout this discussion, students several times displayed confusion about the purpose of the rope; for example, at one point Thomas said, “So the bigger the conflict, the bigger the knot?” Ruth responded: “Well, it's just the way I'm tying them.” At another point, Thomas seemed to believe the rope was representing character change: “So in the beginning of the year, she's red and now, she's like…” Ruth concluded this activity by saying, “Anyway, do you see the conflicts now? Is that getting kind of clear?” Students chorally responded, “Yeah,” and Ruth moved on and let them begin their homework. This second activity with the rope took approximately six minutes.

Ruth’s rationale. Ruth explained that “I knew [the students] had trouble with conflict, and I knew they didn’t understand it.” She chose to use the rope because “they’re such concrete thinkers that they don’t really put things together.” She explained that the rope came from another reading curriculum she had taught, called Project Read. When, after the lesson, she reflected on her use of the rope for the second activity, using
the rope to review all the conflicts that had occurred so far in the text, she said, “I think it made them pause, reflect, think about the different – it was a good review... I think the visual, the actual knotting, was more fun and more engaging for them.” Thus, Ruth chose to use the rope to try to make the concept of conflict more concrete for her students, and perhaps to engage them.

**Responsive to multiple stakeholders.** According to the lesson plan, the teacher was supposed to set a purpose for reading by telling students that events in the story have been building to a climax, and to sum up the lesson by reviewing the events of the story using a narrative arc. The homework assignment was a worksheet about conflict in the text. Therefore, Ruth’s choice to use the rope to teach conflict seemed to be faithful to the intentions of the curriculum developers. It appears that Ruth also considered her students’ need for concrete visuals when designing this adaptation.

**Responsive discourse practices.** Although Ruth attempted to engage her students in discussion around her use of the rope to teach conflict, the resulting discussion was closer to an example of recitation (IRE) than of responsive discourse. Although some content – notably, almost half of the conflicts represented by knots in the rope – was student-generated, Ruth clearly had a series of correct answers in mind, which she attempted to elicit from students through questioning.

**Task complexity.** Ruth’s enactment changed the content of the lesson very little, merely using the rope as an alternative representation of conflict at the beginning of the lesson, and an alternative representation of rising action after reading. Thus, this adaptation could be said to maintain the complexity of the task. However, the confusion that resulted from Ruth’s use of the rope was probably unproductive.
In the next section, I will consider three adaptations made by the second focal teacher, Mary. She was recorded teaching a group of ten sixth graders, four girls and six boys.

**Adaptation 1: Increasing complexity and participation.** The first focal adaptation took place during a lesson on *The Skin I’m In* (Flake, 2000) recorded in January 2016. In this lesson, students were supposed to engage in a debate about who has the power: Char or Maleeka. The formal curriculum referred to this as a “structured debate,” and suggested dividing students into groups of four, with each group assigned to argue for Char or for Maleeka. Within groups, the curriculum suggested assigning each student a role: facilitator, timekeeper, recorder or presenter. Each group would then collaborate to write talking points in support of their argument, using a workbook page as a template. On debate day, presenters would deliver their group’s prepared argument. The lesson plan used language like, “encourage presenters to look up some of the time, not just to read from the page. If they practice enough, they may be able to do this!,” and “The audience may ask questions of each team at the end of their presentation.” Thus, in the formal curriculum, the day’s activity was clearly a “structured debate,” rather than a true debate involving dialogic argumentation (Nielsen, 2013). Mary, on the other hand, engaged her students in a true debate.

The beginning of the lesson plan suggested that in “debate preparation,” presenters from each team should practice delivering their arguments. Instead, Mary began class by giving her students time to gather evidence, continued from a previous day. Students divided into three groups – two for Maleeka, one for Char – and collected evidence from the text on sticky notes. Mary circulated to assist, repeatedly engaging
with students’ thinking, and sometimes facilitating switching of sticky notes among
groups if someone collected evidence that would better support the other point of view.

After about twenty minutes, she asked students to move their desks into two rows
of five, facing each other, representing the Char and Maleeka teams. Introducing the
debate, she said, “I don't want to be in charge of the debate. I want you to be in charge of
the debate and being as grown up as you possibly can.” She selected a student director for
each group to call on their teammates to speak. The debate began with each student
sharing his or her position along with one piece of evidence. Mary intervened three times,
each time to clarify students’ thinking and/or press for reasoning, as in the following
excerpt:

Mary: So why did Maleeka have the power there?
Eric: Because she could have called the cops or walked out.
Mary: Did anyone else know that Char had to do this?
Eric: Only Maleeka.
Mary: So that puts her in power in a way?
Eric: Yeah, because she knows why Char is so bossy.

When all students had shared their positions, Mary opened up the discussion,
asking: “Is there anybody that wants to rebut a statement about Char having power?” For
the remaining 17 minutes of the debate, students engaged in four topical episodes,
defined as “all turns of talk or utterances lying between topic shifts,” of dialogic
argumentation (Boyd & Rubin, 2006, p. 151). The first, lasting four minutes, was about
an episode where Char forced Maleeka to get her a new hamburger, and Maleeka
retaliated by covering the mold on the hamburger bun with ketchup and mustard. Student
Ryan had introduced this topic in the original presentation of his argument as evidence
that Char had the power. In response to Mary’s request for rebuttals, Eric responded: “I
disagree with Ryan because she wasn’t forcing Maleeka to do it. Maleeka could have said
no and walked away. When she moved to Desda, she could have done that when Char was forcing her to but she didn’t. Maleeka had the power to walk away and not do what Char said.” Mary then encouraged dialogic argumentation by returning to Ryan, asking: “Do you agree?” This topical episode continued for four minutes, and involved the participation of two more students besides Eric and Ryan.

The next topical episode was also student-initiated. It began when Mary asked, “Does anyone else have anything they would like to discuss and talk about with another person on the team?” and student Joshua responded, “I have something. Can I do one more of mine?” He then introduced the topic of the deal struck between Maleeka and Char, where Maleeka would do Char’s homework in exchange for borrowing Char’s clothes. This topic was discussed for five minutes.

The two remaining topical episodes were both introduced by Mary: she introduced the third episode by saying, “What about the cigarette incident? Who was in power there?” This topic was discussed for about a minute and a half. Mary then introduced the fourth and final topical episode, Char’s home life, by saying: “So this is what I wonder. Why would [Maleeka] put herself in this situation? What does she know about Char that keeps Char in power? What do we know about her? What do we know about Char? What kind of character is she?” Students discussed this topic for approximately four minutes. Throughout the debate, Mary engaged in talk moves such as pressing for reasoning (“So you’re agreeing with Jack? Why would you agree? How is just putting mustard and ketchup on the hamburger showing that Maleeka has the power?”), encouraging students to say more about their ideas (“So what are you arguing here, that Maleeka could have had the power?”), and inviting other students’ ideas
Near the end of the class period, Mary cut off the debate and asked students to raise their hands to show whether they thought Maleeka or Char had the power. Remarkably, almost all the students raised their hands to indicate that Maleeka had the power—remarkable because, first, this position represented a more inferential interpretation of power in the text, and secondly, because this group of middle school students changed their minds based on a class discussion rather than becoming entrenched in their positions.

**Mary’s rationale.** Mary explained her adaptation to the debate structure:

I didn’t set it up where people weren’t involved. I had tried that where I had teams one and two up here and three and four back here. But it just seemed, for that group in particular, better if everyone was involved and could take, get involved in the discussion. And I think everyone was participating. And that’s what I’m always looking for, 100 percent participation…. [This group of students] is a very immature group. And any time left on their own, they’re just not engaged. I’m looking for engagement and motivation.

Thus, this constituted a “learned adaptation,” where she drew on knowledge from previous enactments of this curriculum to inform her decision (Choppin, 2011). Her goal was to increase student engagement and participation by involving all students in the debate.

**Responsive to multiple stakeholders.** STARI’s theory of change explicitly states the belief that student talk is the mechanism through which deep comprehension will develop. Mary’s rationale for whole-class participation in the debate was to meet her goals of “100 percent participation” and “engagement and motivation,” both goals that are well aligned with STARI’s theory of change. Thus, Mary’s adaptation of this debate to include the participation of all students was clearly faithful to the intentions of the curriculum developers. Furthermore, her adaptation was based on observations of her
students. She had tried using the structure described in the lesson plan before, and felt that it was disengaging to the students who were not actively participating. Thus, this adaptation was responsive to both the curriculum designers and Mary’s students.

**Responsive discourse practices.** Mary’s adaptation to a whole-class debate certainly emphasized the contributions of all students. Her talk moves throughout the debate encouraged students to develop their ideas (as she pressed for reasoning) and to listen to others’ ideas (as she repeatedly invited students to respond to their peers) (Michaels & O'Connor, 2015). Thus, this adaptation supported responsive discourse practices.

**Task complexity.** The lesson plan called for students to read aloud a prepared argument in response to the debate question. Mary required her students to not only to prepare an argument, but to engage in dialogic argumentation - listening and responding to one another in the moment, drawing on textual evidence - a structure that allowed them to modify their own views in light of the evidence presented by their peers. Thus, this adaptation enhanced the task’s complexity.

**Adaptation 2: Adding an activity.** Mary’s second adaptation took place during a lesson recorded in April 2016. At this point in the year, she was teaching Unit 2, which focuses on *Locomotion* (Woodson, 2003). This text, a novel written as a series of poems, is about main character Lonnie’s adjustment to living with a foster family.

Mary chose to begin class with a teacher-created activity in which students wrote original poems, saying, “we said every single day we’re going to be writing poetry, every single day, okay?” She distributed and read aloud a list of phrases that the curriculum had included for a block party to introduce the text (see Appendix B), saying, “If you need
some ideas, I figure that probably all of these...ideas happen in people's lives.” She instructed them twice to write a memory poem, and while students were writing on lined paper, she wrote her own poem on the whiteboard. After approximately seven minutes of writing time, Mary read her poem on the board aloud and then invited students to share their poems (see Appendix B). Three students shared, and Mary responded with quick verbal feedback; for instance, in response to Kevin’s poem about an amusement park visit, she said:

Do you think there's some language you can add to that to make it more dynamic?... How do you feel in your belly? When you go to the top of the rollercoaster – if you were to write another poem, I would like it set when you're at the top, the top, the top, and your stomach is in your mouth, and you feel like you might throw up, and you're screaming and laughing but really you would rather be crying...You know what I'm saying? Just that actual one moment where you were feeling like, “Oh, I’m out of control. My body is going to hurl out of this car, and I'm going to fly across the world.” Add some imagination to it. Someone else want to share?

In response to Sierra’s poem about a car trip to Florida for her birthday, Mary said: “You could have even said on the last line, the best birthday gift ever. Driving hours and hours, 17 hours, oh my gosh. Nonstop? Did you put in the word nonstop?...Someone else want to share?” In total, this activity took approximately 23 minutes.

**Mary’s rationale.** In interviews, Mary spoke at length about her love of writing and her extensive training in writing and the teaching of writing, including work with the National Writing Project. She explained her decision to add poetry writing to the *Locomotion* unit:

I'm very big into writing. Always done a lot of writing with my students. I think it's really important... even for those kids who you would say are basically non-writers... it's that fluidity of being able to understand that you can write...And they can do it, but they need more of that, I think.
Thus, her decision to make this adaptation seemed primarily based on her own strengths and interests as a teacher.

**Responsive to multiple stakeholders.** The lesson plan objectives were for students to “improve reading rate and phrasing,” “notice character traits,” and “notice formatting of poetry as opposed to prose.” STARI is primarily a reading program, which does not include much writing-related instruction. Thus, Mary’s adaptation was not entirely faithful to the intentions of the curriculum developers. Mary believed that writing poetry was beneficial to her students, that it helped them believe that they could write, and that “they love writing the poems.” However, she did not describe any specific observations or data that suggested she had integrated this activity in response to her students’ needs.

**Responsive discourse practices.** Mary attempted to provide feedback on her students’ poems when they read them aloud. However, this feedback tended to take the form of evaluation rather than responsive discourse. In both of the examples provided above, she rephrased sections of the students’ poems, telling them what to add or change, and then moved on to another student without giving Kevin or Sierra the chance to respond. Thus, this adaptation did not reflect reflexive discourse practices.

**Task complexity.** Writing poetry is certainly a complex activity, more complex, many might argue, than reading poetry. Thus, this adaptation did seem to maintain or enhance the overall complexity of the lesson.

**Adaptation 3: Modifying an activity and changing its purpose.** Mary’s third adaptation occurred during the same lesson on *Locomotion*. According to the lesson plan, students were supposed to begin with fluency practice, then read two poems from the text with a partner, and then complete four workbook pages. Mary skipped fluency practice,
the reading and two of the workbook pages; after poetry writing, she spent the rest of the class period focused on one adapted activity from the workbook. The workbook page provided students with all of the words from “Line Break Poem,” the poem on page 4 of *Locomotion*. Students were supposed to cut these phrases out, and then paste them onto the next page of the workbook. The workbook instructed students to “Make it look just like page 4 in *Locomotion*” (see Appendix B).

Instead, Mary instructed students to cut up the lines and paste them onto their paper in any order of their choice, and then to add four more lines, creating an original poem. Students worked on this for approximately twenty-five minutes. Student confusion was evident, as students repeatedly asked questions like: “Can we glue it on the next page? I don’t know where to glue it.” “Can you put them in a different order?” “We’re not allowed to write words on here? They have to be from this?” Mary responded to individuals with clarifications, but it was apparent that her frustration was increasing as she began making comments like, “Kevin, where are your listening skills? I’ve said this four times” and “What’s so hard about this? If you would just sit and concentrate-.” To conclude the activity, five students shared their poems, and then Mary asked, “What are we seeing here with all these lines? Are they all the same?” She attempted to bring the activity back to the text by saying, “‘Line Break Poem.’ Look it up on page four, right there. See how they did it. What’s different between yours and theirs? They’re much smaller, aren’t they? All right. Did you look at the poem on page four? It’s different.” At that point, the class period was over and students began to pack up their belongings.

**Mary’s rationale.** Mary explained that her students gave her the idea for this adaptation, and she decided in the moment to make the change:
A couple of them asked me, “Well, do we have to put it in the same order?”…. And then my creativity took over and I said, “Well, see what you come up with.”… I thought that was a much deeper way of- rather than just copying…. I think we ran out of a little bit of time, but we talked the next day: “Why did the author choose to write it this way?” And that to me is, you're getting into analysis and evaluation.

Thus, she made the adaptation in response to a student’s idea, with the intention of increasing the task’s complexity and making it more “creative.”

**Responsive to multiple stakeholders.** The lesson objective states that students should “notice formatting of poetry as opposed to prose.” It is possible that students could have met this objective through this activity; however, the recorded lesson offered no evidence that this objective was met. Mary’s brief reference to the poem on page 4 of *Locomotion* in the last few moments of the class period was insufficient for students to comprehend the use of line breaks in poetry. Perhaps students met this objective through the next day’s discussion that Mary referenced in her interview; however, based on the lesson recording, this adaptation does not appear to be faithful to the intentions of the curriculum designers.

Mary stated that she made this adaptation in response to her students’ questioning of the activity, and that “they were much more into it” than they would have been if she had implemented the activity as written. Thus, it appears that she made this adaptation in response to a student’s suggestion and in an attempt to make the activity more engaging for her students, though it is unclear whether she considered her students’ skills.

**Responsive discourse practices.** This activity involved very little discussion. Most of the interactions between Mary and the students involved student confusion and teacher clarification. Even when Mary did provide more substantive feedback (e.g., “Think about punctuation as well. Are you going to need quotation marks? Do you want
an exclamation mark? Do you want something said with feeling?”), the feedback tended to be brief and somewhat evaluative, with little room for students to respond. Thus, this adaptation did not reflect responsive discourse practices.

**Task complexity.** Creating an original poem by reorganizing and adding to the language of a poem from the text is certainly more complex than simply “copying” the poem from *Locomotion*. Thus, this adaptation did maintain or increase task complexity.

**Conclusions.** All three of Ruth’s adaptations were faithful to the intentions of the curriculum designers. Furthermore, each of her adaptations had the potential to maintain or enhance the task’s complexity. In her first, most productive adaptation (the character trait activity), she provided students with scaffolds (a trait list and T-chart) to support them in tackling this increased complexity. However, in her other two adaptations, students struggled significantly with the tasks. In addition, only the character trait adaptation reflected responsive discourse practices. In contrast, all three of Mary’s adaptations increased the level of task complexity compared to the activities in the formal curriculum. However, two of the three adaptations were not faithful to the intentions of the curriculum designers, nor did they reflect responsive discourse practices.

After examining all six adaptations, I propose that Debarger et al.’s criteria (2013) for productive adaptations may be multiplicative, rather than additive. In other words, all three criteria must be met in order for the adaptation to be productive. Each teacher made one very productive adaptation (Mary’s debate and Ruth’s character trait activity), in which she stayed faithful to the intentions of the curriculum designers, increased the level of task complexity while meeting her students’ needs with appropriate scaffolds, and engaged in productive talk moves to encourage students to develop their ideas. The other
four adaptations, by failing to meet one of these criteria, must be considered unproductive. For example, Mary’s second and third adaptations, the two that involved student poetry writing, were both complex and creative tasks. However, in neither case was it evident that student learning took place, still less evident that this learning was equivalent (or superior) to that which would have occurred if she had enacted the lesson plan as laid out in the formal curriculum.

In addition, it must be acknowledged that even the two productive adaptations had room for improvement. Although Mary’s interventions into the debate consisted mostly of skillfully deployed productive talk moves, she clearly did not accomplish her own goal of students, rather than herself, being “in charge of the debate.” Furthermore, the debate was dominated by three (male) students. Of the ten in the class, several only spoke when Mary called on them directly. While Ruth’s students appeared mostly to select appropriate descriptors for Momma, she ran out of time before students were able to delve into the more complex part of the lesson, supporting these traits with evidence.

Finally, it is important to note the impact that adaptations have on pacing. Since time is finite, any addition to a lesson plan requires that something else must be omitted. Both teachers were teaching the first novel of the year in January, more than four months into the school year, and the second novel in April, more than eight months into the year (although STARI classes met only three or four days per week). Both teachers complained about difficulties with pacing, saying, “It’s so hard to get through the curriculum…it just takes so long to get through a book,” and “we already spent a huge amount of time…on this book.” However, both teachers routinely cut the amount of text students were supposed to read in a lesson; in both of Ruth’s lessons, she had them read
one instead of two chapters, and in Mary’s second lesson, she skipped the assigned reading altogether. Thus, pacing is a concern when considering adaptations.

**RQ3: Adaptations and Teachers’ Orientations toward Curriculum**

To address my third research question, about teachers’ orientation toward curriculum, I looked across all four teachers’ lesson sets to better understand their responses to the STARI intervention. Figure 3 illustrates each teacher’s two most common specific adaptations. This figure demonstrates that the kinds of adaptations each teacher tended to make were different: for example, Mary tended to fundamentally alter the curriculum by omitting some activities and changing others. While adding new questions was a common adaptation for both Ruth and Christine, Christine did it far more often, adding new questions in 13 lesson segments, as opposed to Ruth’s six. This teacher-level variation demonstrates the importance of examining types of adaptations made by each individual teacher, rather than only looking at the four teachers collectively. Although of course each teacher expressed a variety of complex considerations and motivations for her instructional choices, through the process of creating accounts of practice, I sought coherent representations of each teacher’s orientation toward the STARI curriculum. In this section, I present my accounts of these orientations, supporting each with data drawn from across the lesson set.

**“I’m a rule-follower”: Ruth, a fidelity-oriented teacher.** Ruth was in some ways the most faithful to the formal curriculum – she was the only teacher to follow the guidelines for mode of reading suggested by the curriculum, and she asked students nine of eleven questions from the lesson plans. However, she was also the only teacher besides
Mary to introduce teacher-created activities, spending 25 of 150 recorded minutes on activities that did not originate from the STARI curriculum.

Ruth seemed to make sense of STARI as a program to implement with fidelity. Her instructional emphasis was on following the sequence of activities laid out in the curriculum. She described the purpose of STARI as “to really build reading comprehension….It’s building background knowledge. And it’s the close reading and not just buzzing through something and kind of catching fragments of a passage or book. And deeper thinking.” When queried about the mechanisms through which this occurred, Ruth responded: “Well, I would cite the research for it, which I can't tell you what it is right now.” The theme of the importance of fidelity to the program came up over and over again for Ruth. She explained, “I think that these programs, no matter what it is, are designed by experts. And there's a lot of research that goes into 'em and they're – it's proven research. So I don't tend to adapt too much.” She critiqued Wendy’s choice to skip a unit (described below), saying, “If you look at the progression…I think they need the skills in between. So I probably wouldn’t jump ahead.” Ruth’s instructional emphasis was on getting through the unit as written; four separate times in interviews, she mentioned the importance of getting through the text, with comments like, “I really want to get through this book before vacation,” and “I just want to push and get this finished.”

Yet Ruth was also the most critical of the program of the four teachers, openly rejecting some aspects, particularly phonics activities. Drawing on her expertise as a Wilson Reading (Wilson Language Training Corporation, 2017) teacher, she complained that STARI’s phonics “is not always correct,” and “it’s like it’s a token gesture towards it…. there's not enough of it to really make a difference for poor spellers.” Describing a
conversation with a parent, she told me: “I said, ‘His spelling is not good but we won’t really spend much time on it until after vacation.’ I mean, I hated saying that to the parent.” In other words, she felt constrained by the order in which skills were laid out in the curriculum, such that she did not feel empowered to teach spelling until the point at which it was introduced in the formal curriculum. In response to her frustration with the phonics instruction in the curriculum, she skipped most of it. She was critical of other aspects of the curriculum, too, for example, workbook pages where students were asked to guess the meaning of unfamiliar words. She protested, “I think that’s really bad to do that...because that’s what sticks in their head...I just think this is a really bad page and should not be used. They’re not using the words; they’re making up meaning. They’re not looking at a base word to try to figure out what this could possibly mean.” She also critiqued the pacing of the curriculum, saying, “it's so hard to get through the curriculum. If there were one criticism I would have, it's that the units [are too] long.” However, she also reduced the amount of text students read as compared with the amount in the lesson plan in two of her three recorded lessons (and in the third, the total text in the lesson was only three poems).

Furthermore, despite Ruth’s characterization of herself as a rule follower who implemented the program with fidelity, unlike Wendy and Christine, she actually made significant substantive adaptations. Mostly, the activities she introduced were drawn from other programs; for example, she used a rope made of colored yarn to teach conflict, drawn from a program called Project Read. She qualified her use of this adaptation by admitting her lack of fidelity to the other program: “I mean, really to do this, I would have done the whole preliminary thing from another program, but I just needed to do
something quick because they didn’t get it.” Thus, Ruth’s enactment of the curriculum actually seemed to conflict with her orientation towards it, though she did not seem to realize this. (In fact, in her final interview, when I queried her about specific adaptations, she responded, “Didn’t think I made that many changes.”)

“The books are really so interesting”: Christine, a content-oriented teacher.

Christine made few major adaptations to the formal curriculum, though she varied the style of implementation. She chose to use round robin reading rather than the modes of reading suggested in the lesson plans, added numerous discussion questions and addressed additional vocabulary besides that selected in the plans. She also spent significant time on fluency, largely because she incorporated discussion into fluency practice.

Christine seemed to make sense of STARI by focusing on the content of the curriculum; her words and actions indicated that student comprehension of the content of the text was her most important goal. In service of this goal, Christine added supplementary vocabulary instruction: she created a word wall where she posted all the words listed in the curriculum, and she also paused during reading in the recorded lessons to teach eleven additional vocabulary words. Nine of these were rare words (i.e., words that are not considered general service (Browne, Bauman, & Culligan, 2013) or academic words (Coxhead, 2000)), and not always particularly essential to comprehending the text, as Christine herself acknowledged:

Well, once we talked about [Vicks rub], some said, “Oh, yeah, yeah,” especially the blue jar and it smells like eucalyptus – or menthol – they don't know what that smell is. They were thinking mint. Not that it's a big part of the book, but still I find a lot of times I'm spending more time building general knowledge that you'd think maybe they might have.
Christine seemed to see the primary purpose of the fluency passages as building background knowledge to aid in comprehension. To that end, she adapted the structure of fluency practice so that students were paired with different-level, rather than same-level partners, so that every student would be exposed to two different passages. She also added time for students to share the content of each passage with the whole group. She explained, “We always talk about the fluency passages so that they all can get background information from each other about whatever the topic is. Which I think has been really helpful.” In the last unit of the school year, on *Ninth Ward* (Rhodes, 2010), Christine decided not to include fluency practice in order to save time so that her students could “get through the novel” before the end of the year. When I asked her about this decision, she explained: “I thought, ‘Well, they don't really need to know a lot about Mardi Gras or – I could fill them in on that as we read.’ So that's sort of what I decided to do – to take the pieces that were related to the culture and fill them in as we go.” Building fluency skills seemed to be almost an afterthought; when I followed up by questioning whether she thought her students’ skills were sufficient to forgo fluency practice in the final unit, Christine responded: “Yeah, I'd say so. I mean, they're phrased, they're fluent.” Thus, even in fluency practice, Christine appeared to be more focused on building knowledge to aid in the comprehension of content than in building reading skills.

In addition to a focus on the content of the text, Christine was very focused on reading her students. Instead of engaging in guided reading or partner reading, as the formal curriculum directed, in all three recorded lessons, Christine had her class sit in a circle and engage in round robin reading of a section of text. She explained this choice in two ways. First, she referenced the range of students’ skills:
I very rarely have them do silent reading, because their reading rates are so different that the kids who are really, really struggling and slower readers are always watching everybody else to see what they're doing, and they're not concentrating, and they don't know what they've read... The good readers can answer the questions and participate and those that haven't really read – who've pretended to read - don't participate.

Secondly, she felt that round robin reading allowed her to more effectively build students’ understanding of content; as she said, “I'll watch them as they're reading, and I see like they're not really getting it.” For example, in one recorded lesson, a student stumbled when reading the word “institution,” and Christine stopped to teach the class the meaning of the term.

Christine also expressed concern about the affective engagement of her students. She saw class discussion both as a way to build comprehension of the text (a reason she mentioned three times) and to engage students (which she mentioned four times). In every interview, she also talked about the role of STARI in building her students’ confidence, saying, for instance: “[Students] say this is like a family. It’s cute. It’s nice. They feel comfortable and it’s because they are in small groups and it affords them the opportunity to speak their mind and ask questions that really muddle through things without feeling embarrassed.”

“Your performance is when you’re discussing it:” Wendy, a reading-oriented teacher. Like Christine, Wendy made few major adaptations. She, too, chose to use different modes of reading than those suggested in the lesson plan, and spent significant time reviewing work from the previous lesson. Wendy described the purpose of STARI as being “to increase their reading fluency, their comprehension, and get them right back on grade level. So if they were just a little below, we'd kind of boost them up a little bit.” She believed this would happen through two mechanisms: the extra time
students spent reading as a result of being enrolled in STARI, and discussion. Wendy’s belief in the importance of time on text predated her exposure to STARI, but she twice explicitly stated that she learned about the importance of discussion from her STARI coach. She explained:

[The coach’s] feedback was helpful because she would make sure I understood the idea that the talk is the important part. Because I think at first, you think, “Well, they need to read, read, read.” But it's really the discussing what they read that deepens their comprehension. So sometimes I feel like I'm cheating by just sitting there talking with them. But she made me understand that that was really the important part.

Wendy’s lessons reflected this, as she devoted significant class time to discussion – about 25 of 50 minutes, on average - and she engaged in productive talk moves like pressing for reasoning (Michaels & O'Connor, 2015).

Wendy’s belief in the importance of reading was based in her personal biography. She explained:

when I got to this age, I didn't read anymore. And I was an A student! But I figured out – I'd just sit there and listen. I mean, you don't have to read anymore. So you never got better. For me, I would rather listen and write down my answers, no homework. I'd go out and play. …[So] I know when I was an ELA teacher and assigned them reading, they didn't do it. If you don't give 'em, in school, time to read, they're not going to do it.

Wendy made several adaptations to the curriculum that she attributed to her belief in the importance of time on text. First, she skipped units two and three in the curriculum and went straight to unit four because she felt that units two and three didn’t require enough reading. She explained that she skipped unit two, which focused on poetry, because “when you have a poem it takes you a while to think about it and analyze it, so I was thinking they're not getting much reading.” Secondly, instead of alternating between guided reading and partner reading, she either assigned her students to read independently, or gave them the choice of reading independently or with a partner, on
alternating days. She explained, “otherwise some kids might never do that silent in-your-
head reading.” Instead of conducting guided reading lessons, where she stopped students
every few pages to ask a discussion question, she provided them with all the discussion
questions ahead of time, on the board or on a handout, and asked them to make notes for
a discussion after they finished reading the text. Her reason for this was that, similar to
Christine, she had encountered a problem with the wide range of reading rates in her
class:

at least this way, that kid who is done can start gathering their ideas for this,
where the kid who isn't done might not have time to jot down their ideas, but
they'll still have time to get the reading done…. When they're sitting there and
they're just waiting for their class – and the classmates who aren't done, they're
gonna figure out fast, “Just pretend you read it and don't read it.”

Finally, Wendy’s flexible sense of pacing created additional adaptations. In all
three of her recorded lessons, the bell rang while students were in the middle of an
activity, and she responded with some version of: “We’ll have to do it tomorrow.” She
never closed a lesson with the “sum up” activity described in the formal curriculum.
Instead, she began each lesson by reviewing work from the previous lesson, spending
between twelve and twenty minutes reviewing. Wendy’s flexibility around pacing
predated her work as a STARI teacher; she explained: “In fact, some science classes,
some groups would be so far behind that I would just eliminate an activity or two just to
try to get them caught up.”

“And then my creativity took over:” Mary, a creativity-oriented teacher.
Mary seemed to use the formal curriculum more as a jumping off point than as a script,
frequently changing the use and purpose of lesson materials. She spent about 40% of
recorded class time (55 of 145 minutes) on teacher-created activities that did not originate
in the STARI curriculum, such as creative poetry writing. In the three recorded lessons,
she entirely skipped such curriculum activities as fluency practice, workbook pages and homework.

According to Mary, the purpose of STARI is “to help those students that are still struggling as readers to become much more effective.” She believed that the mechanisms through which this would take place were phonics instruction and the four reciprocal teaching strategies. She emphasized the way in which these strategies spiraled throughout the curriculum and were repeated again and again. She also mentioned the importance of STARI being “an adolescent program so the books are designed with a high interest level.”

Evident in Mary’s orientation toward STARI were her sense of herself as a creative person, and her experience as a gifted and talented teacher. She explained: “I was a gifted and talented teacher for ten years. And I know that affects how I teach, definitely. It makes a difference.” Mary adapted the curriculum more freely than any of the other teachers; in fact, she described her planning and preparation process as a process of adaptation: “you've got to read it over, and know what you're going to do, what pages you've got to include, what pages are you not going to include, what are you going to do instead? That type of thing. And if I need activities, then I'll plan those, as well.” She added significant writing to the curriculum, based on her own extensive background in writing, including training with the National Writing Project and earning credits toward a graduate degree in professional writing.

Mary, however, clearly stated a belief that adaptations were more appropriate for higher level students. She explained: “I just feel they needed more sophisticated. They didn't need the word work and the phonics type of work as much. I'm looking for the
analysis. I'm looking for the higher-level thinking, especially in groups that I feel can handle it.” In contrast, with the groups of students that Mary believed were lower-level, “we stick directly with this, so I'm relying on that manual to carry me through completely in total. Because I think that they need this structure of follow along page by page by page.” While teaching STARI, she explained, her viewpoint on lower-level work changed:

I also noticed like on pages in here where they have you cut up and paste, and I would always say, "Oh, that's too time consuming." But I really found out that when they cut and they paste…that's when they really start thinking about what they're doing. And when they start questioning, "Does it belong here?" They were doing double vowels and saying, "Which list would it go in?"

However, in one of Mary’s recorded lessons, she taught a group she believed to be “higher-level,” while in the other two she taught a “lower-level” group, and she made no more adaptations with the higher-level group. In reflecting on an adaptation she made with her lower-level group where she essentially had them create a poem rather than copy one from the text, she explained:

You're getting into analysis and evaluation. So that's always with me at the top of everything I do is the Bloom’s, the analysis, the higher-level thinking skills. And I know that's, as I said, it's my gifted and talented that comes in because you're always doing that with higher level [students]…But it's important for these kids, too, I think.

In other words, Mary seemed to believe that the purpose of adaptations was to increase the complexity of the formal curriculum, which made adaptations more appropriate for higher-level groups. However, in this instance she stated a belief in the importance of lower-level students also having access to the analysis and evaluation skills she believed she was introducing with this adaptation.

**Orientation and response.** In this section, I will consider each teacher’s orientation toward STARI in light of Coburn’s (2004) framework for responses to policy
initiatives. While none of the teachers demonstrated rejection or symbolic response, the
four teachers displayed both assimilation and accommodation.

**Assimilation vs. accommodation.** Ruth, the “rule follower,” took a procedural
orientation toward the STARI curriculum, viewing it as a series of steps to be followed
with fidelity. Thus, in Coburn’s framework, Ruth demonstrated assimilation, meaning
that she “understood messages in terms of changes in instructional routines, materials, or
classroom organization, rather than in terms of the underlying pedagogical or
epistemological assumptions of the approach” (Coburn, 2004, pp. 224-225). This is not to
imply that Ruth was ineffective or mechanical in her implementation of STARI; she was
typically responsive to students, pressing for reasoning when they attempted to answer
her questions and sometimes bringing in new activities to engage them or to scaffold their
understanding of concepts. Ruth also spoke highly of the curriculum, and of the
professional development she had received. Thus, Ruth’s assimilation of the curriculum
does not imply a misunderstanding of the pedagogical assumptions of the curriculum –
merely that her orientation as a self-described “rule follower” outweighed this
understanding in her enactment of the curriculum. Ironically, however, this orientation
did not mean that her enactment of the curriculum was more aligned with the formal
curriculum than the enactments of other teachers. It did mean that her critiques and areas
of disagreement with the curriculum seemed to cause her more frustration than they
caus ed the other teachers, perhaps because she felt less empowered to make adaptations.

Christine’s enactment of the STARI curriculum reflected her orientation toward
text and toward students. She believed in the importance of students’ mastering the
content of the text, and to that end she taught a lot of additional vocabulary words, even
words that were not, in her own estimation, “a big part of the book.” Her use of the units’ fluency passages was geared more toward building background knowledge than toward building skills. She modified the mode of reading, using round robin reading so that she could more closely monitor her students’ comprehension of the text and jump in with additional instruction at any moment – an adaptation that she saw as responsive to her students’ needs. Furthermore, she believed that STARI played a role in building her students’ confidence and affective engagement with text.

Thus, Christine’s enactment of STARI also represents assimilation in that she “interpreted and enacted messages in ways that transformed them to fit [her] underlying assumptions” (Coburn, 2004, p. 224). Her prior belief, based on more than twenty years as an elementary school teacher followed by five years as a middle school reading interventionist, was that the role of a reading teacher was to support students in comprehending text. Thus, Christine focused on close reading of the text and close reading of her students in order to scaffold their understanding of the text. She devoted significant time to students’ comprehension of even less-than-essential aspects of the text, like Vicks rub. Skills instruction played a less important role in Christine’s enactment of the curriculum.

Wendy’s enactment of the STARI curriculum reflected her belief in the importance of time on text, which caused her to adapt both the units she chose to teach and the mode of reading that occurred during those units. However, through work with her STARI coach, she gained a belief in the importance of discussion, which prompted her to engage students in frequent high-level talk around text. Thus, Wendy demonstrated both assimilation and accommodation of STARI’s messages. The changes she made in
service of her belief in students’ time on text – substituting independent reading for
guided reading, and changing the order of the units she taught – represented assimilation,
or transforming the curriculum to fit her preexisting beliefs. Furthermore, her flexible
pacing – frequently ending instruction in the middle of an activity when the bell rang,
seemingly unexpectedly, and then beginning the next class with an extended review of
the previous lesson – also reflected assimilation, as she had taught this way prior to
beginning to teach STARI. She believed that she was being responsive to students with
this flexibility; however, this practice meant missed chances to consolidate student
learning through the daily “sum up” activities, and the extended reviews also slowed her
progress through the unit.

In contrast, Wendy’s belief in the importance of discussion, which she explicitly
stated that she learned from her STARI coach, demonstrated accommodation, or
“restructur[ing] [her] fundamental assumptions about the nature of reading instruction or
students’ learning” (Coburn, 2004, p. 225). Wendy’s words, “I feel like I’m cheating by
just sitting there talking with them,” reflect what a fundamental restructuring of beliefs
this was for her. As in Coburn and Woulfin’s (2012) research, which found that teachers
were more likely to respond with accommodation when a policy message came from a
coach, Wendy’s work with her STARI coach caused her to change her approach to
reading instruction.

Mary made sense of the STARI curriculum through her belief in the importance
of creativity and her experience as a gifted and talented teacher. She felt free to adapt the
curriculum, and made more substantive adaptations than any other teacher. She stated a
firm belief that adaptations were more appropriate for higher level students, but in this
lesson set was equally likely to adapt the curriculum regardless of her perceptions of the
ability level of the group. Thus, Mary also responded to STARI with assimilation,
aligning its practices with her preexisting beliefs and experiences. She demonstrated
some accommodation of the program’s philosophy for lower-level learners, for instance,
recognizing the value of cutting and pasting in learning spelling patterns. However, she
believed that more substantive adaptations were for higher-level groups who “needed
more sophisticated.” It appeared, though, that she made just as many adaptations with her
“lower level” groups, and explained, “it’s important for these kids too, I think.”
Furthermore, many of the adaptations she made in order to increase complexity and build
creativity – such as turning an exercise in understanding line breaks into a creative
writing opportunity – were not necessarily aligned with the intentions of the formal
curriculum. Thus, Mary transformed most curriculum activities to fit with her preexisting
worldview.

**Parallel structures.** Coburn’s (2004) research took place in self-contained
elementary school classrooms, and her example of parallel structures involved teachers
engaging in two conflicting practices of reading instruction with the same group of
children. The institutional structure of contemporary American middle schools, where
some teachers teach different subjects to different groups of children all day, may
inherently breed parallel structures. All four STARI teachers taught other classes in
addition to STARI – Ruth taught Wilson Reading, and Christine, Wendy and Mary taught
LLI. A final factor in a complete consideration of the teachers’ orientations toward
STARI is the extent to which they practiced parallel structures in their instruction.
Neither Ruth nor Christine saw any comparison between their STARI instruction and
their other classes; in Ruth’s words, “It’s really apples and oranges.” When I asked Christine why she thought her LLI instruction was dissimilar to her STARI instruction, she replied: “I think everything [in LLI] is more black and white, it's right there...This [STARI] story has more inference, the language is rich.” Both of these responses seemed well-aligned with the two teachers’ orientations toward curriculum. Since Ruth viewed both instructional programs as (different) series of steps to be followed, she saw no resemblance between them. Since Christine made sense of both programs through the content of the text, she saw little similarity, because the STARI texts were richer.

When I asked Wendy, on the other hand, if she had changed her teaching of LLI in any way since beginning to teach STARI, she replied: “Maybe I do have more conversation about the books, too, especially before they write. We definitely talk a lot. So maybe that did carry over and I didn't realize it.” This is further evidence for Wendy’s accommodation to the importance of discussion – her belief in this pedagogical principle extended into her non-STARI instruction.

Mary, however, took perhaps the strongest stance against parallel structures: midyear, she transitioned all of her classes to the STARI curriculum. She explained that she was “talking to someone who said that LLI was really more for younger elementary. And I thought, ‘You know, I really think that's true, that you have a much broader skill [in STARI].’” Thus, this choice seems aligned with Mary’s orientation toward pushing more complex tasks. Furthermore, she explained that “with LLI I was planning things all the time. That was much more burdensome,” whereas with STARI, “it's just right there on the pages, and you can step right in and you can teach it, and you don't have to be making up activities and trying to figure out where the kids are going to – it's all guided.”
Therefore, like the teachers in other research (e.g., Datnow & Castellano, 2000), Mary may also have made this change to make her workload more manageable.

**Different orientations lead to varied enactments.** The four teachers’ different orientations toward the curriculum led to adaptations that resulted in four quite different enactments. In a number of cases, multiple teachers shared the same concern or critique, but responded in different ways. For example, Wendy and Christine were both concerned that there was not enough reading in the poetry-focused unit, but only Wendy made a change and substituted a different unit. Christine, Mary and Ruth all expressed concern about the lack of writing in the STARI curriculum, and felt that they should add more writing instruction; however, only Mary added writing to her lessons. From both a research and practice perspective, it is important to understand the variation with which these four teachers enacted the curriculum.

Furthermore, it is critically important to understand the ways in which teachers’ orientations toward curriculum shape their enactments. Coburn (2004) asserts that teachers’ responses to policy messages depend on the degree of congruence the messages have with their preexisting beliefs and practices. None of the teachers’ preexisting orientations – Ruth’s belief in the importance of fidelity, Christine’s belief in the importance of text, Wendy’s belief in the importance of reading, and Mary’s belief in the importance of complexity and creativity – could be considered incongruent with the principles of STARI. However, teachers’ processes of making sense of STARI through these orientations sometimes produced *practices* incongruent with those principles. For example, both Wendy and Christine encountered the same problem with guided reading due students’ varied reading rates. Based on her belief in the importance of reading,
Wendy’s response was to provide her students with a written copy of the discussion questions ahead of time, so that students who finished early could use the extra time to make notes to prepare for discussion. This adaptation could be considered a “tweak,” still aligned with the intentions of curriculum designers. Christine, however, made sense of this dilemma through her belief in the importance of students comprehending the content of texts and chose to engage in exclusively round robin reading. Current research suggests a number of problems with round robin reading; for instance, it may contribute to students developing ineffective reading habits by forcing them to slow their silent reading rate to match a peer’s oral reading rate; it may result in the reader’s miscues being corrected too quickly without giving the student a chance to self-monitor, and it may hamper listening comprehension if students rely more on the visual than the auditory representation of the text when exposed to both simultaneously (Opitz & Rasinski, 2008). Thus, this adaptation might be considered a “lethal mutation,” one which departs dramatically from the principles of the formal curriculum (Brown & Campione, 1996, p. 292).

**Implications**

The present study analyzed the adaptations made by four experienced teachers engaged in sustained implementation of an adolescent reading intervention, after the formal randomized controlled trial had concluded. Through observation and interview, I identified the adaptations each teacher made to the curriculum. I evaluated the productivity of six focal adaptations, and then analyzed what each teacher’s adaptations suggested about her orientation toward the curriculum. Given the lack of attention to sustained intervention implementation in the existing literature, this study suggests
several important implications for development, implementation and analysis of future reading interventions.

Findings from the present study show, first, that it is possible for a school to sustain implementation of an intervention curriculum after the intervention study has concluded. Although only one of the nine original treatment schools formally sustained implementation, the four teachers in that school taught STARI every day. However, the teachers adapted the curriculum so much that the experiences of students in each teacher’s classroom were quite different from one another, and in some cases markedly different from the intentions of the curriculum designers.

Furthermore, despite the extensive professional development teachers received in implementing this curriculum, only one third of the six focal adaptations were productive. To some extent, this reflects the challenges inherent in teaching adolescents reading significantly below grade level. Often, teachers’ rationales for their adaptations involved trying to engage a population of students who are often disengaged, or attempting to differentiate to meet the wide range of needs their students presented. However, these attempts were not always successful.

Perhaps this finding should lead to a call for a renewed emphasis on fidelity; however, previous research suggests that an emphasis on fidelity is ineffective. SFA, for example, has, over the course of more than 20 years, shifted away from a fidelity-focused approach to implementation:

because it found that schools and trainers were overly focused on visible details of the implementation rather than the theory underlying their use. Superficial engagement with the program seemed to result in compliance to mandates and fidelity to implementation but did not necessarily translate into enhanced student outcomes...Thus, SFA increasingly emphasizes an understanding of the theory behind the [curriculum] tools, as it has found that it was important for teachers to
understand the concepts behind the tool in order to utilize the tools more effectively or adapt them to fit the needs of students (Datnow & Park, 2010, p. 85).

Drawing on lessons from SFA and other programs, I suggest that rather than an increased focus on fidelity, curriculum design and teacher training should be oriented toward supporting teachers in making productive adaptations even in early stages of implementation. The research presented in Chapter 2 of this dissertation, which shows substantial teacher-level variation in fidelity of implementation during the randomized controlled trial phase of implementation, suggests that teachers make significant adaptations even when instructed to implement with fidelity. Therefore, since adaptation is inevitable, professional development at all stages of intervention implementation should focus on increasing the productivity of these adaptations. Teachers should be taught the three criteria for productive adaptations described in this paper, and asked to use these criteria to examine their own adaptations.

In addition, drawing on recent work on educative curriculum materials in math and science, I propose that curriculum materials should provide teachers with rationales for both content and instructional strategies (Ball & Cohen, 1996; Davis & Krajcik, 2005; Davis et al., 2014; Stein & Kim, 2009). According to Remillard (2000), curriculum materials typically “speak through teachers, by guiding their actions. [They do] not speak to them about these tasks or the ideas underlying them” (p. 347). Thus, teachers are positioned as passive “transmitters” of curriculum (Ben-Peretz, 1990), with their “implicit role … to ‘deliver’ activities to students” (Remillard, 2000, p. 346). However, research suggests the importance of teachers understanding the “first principles” of a reform in order to adapt it successfully; if they do not deeply understand these principles, their adaptations are unlikely to be aligned with the intentions of the curriculum.
developers (Han & Weiss, 2005; McLaughlin & Mitra, 2001). This also underscores the importance of sustained implementation of interventions, as teachers are unlikely to reach this understanding of deep principles in early enactments of a new curriculum.

Finally, it is interesting to note that even within the same school context, teachers’ enactments of the curriculum were quite different. This points to the importance of a focus on teachers’ orientations toward curriculum. A better understanding of teachers’ orientations and sense-making processes may allow curriculum developers to provide more tailored professional development to help teachers understand the ways in which they may adapt a curriculum to fit their worldviews, while still being faithful to the intentions of the curriculum designers. Only through a fuller understanding of teachers’ sense-making processes will future research and practice be able to encourage accommodation to the most critical pedagogical features of a curriculum.

Limitations and Conclusion

In conclusion, I would like to recognize the challenges inherent in implementing a complex curriculum, or in working with students significantly below grade level, let alone doing both at the same time. Therefore, like Charalambous and Hill (2012), I ask that the work presented above “be read more as an analysis of the work of teaching rather than as a critique of the teachers” (p. 461; emphasis in original). In fact, I am tremendously grateful to the four teachers who allowed me to visit their classrooms, and used their limited time to share their perspectives with me.

While the narrow focus of this study – four teachers in a single school - presents obvious limits to generalizability, the present student offers an important contribution to the literature, as previous research on teacher adaptations to literacy curricula is so limited. This study provides valuable insight for future developers of intervention...
curricula by helping us understand the ways in which teachers productively and unproductively adapt curricula. Furthermore, this study examines sustained implementation of a literacy curriculum, an area which is rarely researched.

The present study supports the findings of previous researchers that suggest that while adaptation is inevitable (Burkhauser & Lesaux, 2015; Datnow & Castellano, 2000; Leko et al., 2015), productive adaptation is difficult (Ben-Peretz, 1990; Taylor, 2013). Therefore, curriculum developers and coaches should plan with adaptation in mind. Curriculum should be designed to speak “to” teachers by including educative materials. In addition, the professional development that takes place with the introduction of any new curriculum should include training, coaching and practice in productively adapting the curriculum to meet the needs of a teacher’s unique situation. Curriculum cannot meet the needs of struggling learners while bypassing, or speaking “through,” teachers, as teachers always adapt curriculum. Therefore, only by supporting teachers in making productive adaptations can we hope to improve learning outcomes for students.
Table 1

Demographic Characteristics of Students

<table>
<thead>
<tr>
<th>Measures</th>
<th>% of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Special Education</td>
<td>18.7%</td>
</tr>
<tr>
<td>Low Income</td>
<td>35.4%</td>
</tr>
<tr>
<td>English Language Learner</td>
<td>7.4%</td>
</tr>
<tr>
<td>European American</td>
<td>69%</td>
</tr>
<tr>
<td>African American</td>
<td>10.4%</td>
</tr>
<tr>
<td>Latino</td>
<td>10.2%</td>
</tr>
<tr>
<td>Asian</td>
<td>1.5%</td>
</tr>
<tr>
<td>Native American/Pacific Islander</td>
<td>3.5%</td>
</tr>
<tr>
<td>Mixed/Other</td>
<td>5.4%</td>
</tr>
<tr>
<td>Total Enrollment</td>
<td>461</td>
</tr>
<tr>
<td>Parent Code</td>
<td>Child Code</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Addition</td>
<td>Activity</td>
</tr>
<tr>
<td></td>
<td>Follow-up question</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>New question</td>
</tr>
<tr>
<td></td>
<td>Vocabulary instruction</td>
</tr>
<tr>
<td>Deletion</td>
<td>Homework assigned</td>
</tr>
<tr>
<td>Modification</td>
<td>Different use of materials</td>
</tr>
<tr>
<td></td>
<td>Mode of reading</td>
</tr>
<tr>
<td>Unadapted</td>
<td>LP question</td>
</tr>
</tbody>
</table>
Figure 1. Sample STARI lesson plan
Figure 2. Total adaptations by type and teacher

Figure 3. Most common adaptations by teacher
Chapter 4.

Conclusion

Taken together, the two studies of this dissertation demonstrate that substantively meaningful teacher-level variation occurred in implementation of the Strategic Adolescent Reading Intervention. The two studies present this finding in different contexts, which are important for different reasons. The first study shows that significant teacher-level variation occurred even within the context of a randomized controlled trial, in which teachers understood that they were participating in a research study and that they were expected to implement the curriculum with fidelity. The second study demonstrates that meaningful variation occurred even among teachers within the same school, for whom most contextual variables – administration, schedules, school culture and student population – were the same. Thus, this variation appeared to depend on factors within the teachers themselves.

The second study also demonstrates that teachers’ adaptations to the curriculum were more often unproductive than productive. However, adaptations are inevitable (Burkhauser & Lesaux, 2015; Datnow & Castellano, 2000; Leko et al., 2015), and insisting that teachers implement curricula with script-like levels of fidelity is ineffective (Datnow & Park, 2010). Therefore, in order for literacy interventions to be effective, teachers need to be supported in making productive adaptations to the formal curriculum. Professional development should help teachers to develop deep understanding of first principles of an intervention (Han & Weiss, 2005; McLaughlin & Mitra, 2001). Teachers’ preexisting orientations toward curriculum should be acknowledged and interrogated, in the hopes of guiding them towards accommodation, rather than
assimilation, of an intervention’s most critical pedagogical features. Sustained implementation of interventions will be critical in enabling this, as this kind of deep understanding must develop over time, and it is unrealistic to expect teachers to demonstrate it during their first interaction with a new curriculum.

Secondly, intervention research must take teacher-level variation into account in analyzing the effectiveness of an intervention. It is not enough to cite overall high average levels of implementation to claim that student outcomes are a valid assessment of a program as written. An accurate and meaningful assessment of the impact of a reading intervention must document the process of implementation, and account for teacher-level variation in statistical models evaluating the effectiveness of the program. Considering these factors will allow us to better understand what works, for whom and under what conditions.
Appendix A – FoI protocol, with subscales operationalized by dimension of fidelity

**STARI OBSERVATION RUBRIC revised December 2013**

**Section A: General Lesson Information**

A1. Date: ________________________  A2. Observer: ________________________________


A5. Number of Students Present: _____  A6. Number of Students Absent: _____

Lesson Observed:
A7. Level: 1  2 (circle one)  A8. Unit # _____  A9. Day # _____

A10. STARI Lesson Start Time: ______________  A11. STARI Lesson End Time: ______________

A12. Total STARI Lesson Time: _____________ min  A13. Grade Level: _______________

---------------------------------------------------------------

**Section B: Lesson Part I: REVIEW HOMEWORK**

2  0
B1.  ■ Happened  ■ Did not happen

B2. Teacher checks that individuals have completed their homework

2  1  0
■ Present  ■ Partly present  ■ Not present

B3. Teacher calls on students to share homework responses

2  0
■ Present  ■ Not present

---------------------------------------------------------------

**Section C: Lesson Part II: FLUENCY WORK**

2  0
C1.  ■ Happened  ■ Did not happen

C5. Teacher circulates during fluency work and offers support with the process

☐ continually  ☐ occasionally  ☐ not present

C6. Students are grouped into partners for fluency work

2  1  0

☐ all or nearly all students  ☐ some students  ☐ no one

C7. Both partners have a chance to read a passage aloud during fluency work

2  1  0

☐ all or nearly all students  ☐ some students  ☐ no one

C8. Students record elapsed time and words per minute (WPM) during fluency work

2  1  0

☐ all or nearly all students  ☐ some students  ☐ no one

C9. Students in the class are working in more than one fluency level, A-D

2  0

☐ present  ☐ not present

C10. Students read phrase-cued passage or challenging phrases and words out loud during fluency

2  1  0

☐ all or nearly all students  ☐ some students  ☐ no one

C11. Students record answers to comprehension questions about fluency passage

2  1  0

☐ all or nearly all students  ☐ some students  ☐ no one

C12. Students ask each other and answer follow-up questions about the fluency passage or comment to their partner about the passage content

2  1  0

☐ present  ☐ partly present  ☐ not present

---------------------------------------------------------------------------------------------------------------------

Section D: Lesson Part IIIA: Guided Reading (Days 16-30)
Note: For days 16-30, the class is split into two groups during Part III. Please observe one guided reading group. While the other guided reading group meets, please observe partner work.

D1. How many guided reading groups are there?
   1       2
   □ one       □ two

D2. □ Both groups met □ One group met □ No groups met


D6. Total Time: _____ min
D7. Time with Group 1: _____ min  D8. Time with Group 2: _____ min

# of students in guided reading groups:  D9. Group 1: _________  D10. Group 2: _________

STUDENT ACTIVITY - BEFORE GUIDED READING

D11. Students sit in a group with the teacher with copies of the guided reading book
   2       1       0
   □ all or nearly all students □ some students □ no one

D12. Students are seated so that they face each other and the teacher (e.g. in a circle or square)
   2       0
   □ present □ not present

D13. Students have materials (post-its or bookmarks or notebooks) to record new words and/or mark quotes during guided reading
   2       1       0
   □ all or nearly all students □ some students □ no one

D14. Students participate in summarizing the previous day’s guided reading passage
   2       1       0
   □ present □ partly present □ not present

D15. Students participate in discussing the new words for the guided reading passage they will read next
   2       0
TEACHER ACTIVITY– BEFORE GUIDED READING

D16. Teacher leads a summary discussion of the preceding day's guided reading passage reading with the group

<table>
<thead>
<tr>
<th>present</th>
<th>partly present</th>
<th>not present</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

D17. Teacher uses a whiteboard or projector to introduce new words before reading

<table>
<thead>
<tr>
<th>present</th>
<th>not present</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

D18. Teacher talks about the new words in a meaningful context

<table>
<thead>
<tr>
<th>all words</th>
<th>some words</th>
<th>none of words</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

D19. Teacher sets a purpose for reading the next section of the novel

<table>
<thead>
<tr>
<th>present</th>
<th>not present</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

STUDENT ACTIVITY – DURING AND AFTER GUIDED READING

D20. Students read silently as directed

<table>
<thead>
<tr>
<th>all or nearly all students</th>
<th>some students</th>
<th>no one</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

D21. IF “no one”, other approach:  

<table>
<thead>
<tr>
<th>students read aloud</th>
<th>teacher reads aloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

D22. Students participate in discussion of guided reading novel

<table>
<thead>
<tr>
<th>all or nearly all students</th>
<th>some students</th>
<th>no one</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

D23. Students provide extended responses (more than a short phrase answer or comment) during discussion of the novel

<table>
<thead>
<tr>
<th>present</th>
<th>partly present</th>
<th>not present</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
☐ frequently  ☐ once in a while  ☐ never

D24. In discussion of passage meaning, students reference text explicitly (e.g., by pointing to a paragraph or reading a quote)

☐ frequently  ☐ once in a while  ☐ never

Tally all text references made by students: _________________________ D25. Total #: ____

Students make connections to prior knowledge or experiences
Tally all such connections made by students: _________________________ D26. Total #: ____

Students explicitly connect their contributions to each other (e.g., “I agree with...”)
Tally all peer connections made by students: _________________________ D27. Total #: ____

TEACHER ACTIVITY - DURING AND AFTER GUIDED READING

D28. Teacher directs students to silently read particular text chunks and then stop for discussion

☐ present  ☐ not present

D29. IF other approach:

☐ students are directed to read a whole chapter or extended section  1

☐ students are directed to read out loud rather than silently  2

D30. When directing students to silently read a chunk of the novel, teacher provides a context or a purpose for reading that chunk of text

☐ present  ☐ partly present  ☐ not present

D31. Teacher poses literal (“right there”) questions:

☐ present  ☐ not present

D32. Teacher poses “search and think” questions (e.g. how, why, describe, explain, compare):

☐ present  ☐ not present
D33. Teacher asks students to re-read or refer back to text:

2  0

☐ present  ☐ not present

D34. Teacher asks follow-up questions to elicit fuller or clearer student responses (e.g. “say more about that”):

2  0

☐ present  ☐ not present

D35. Teacher explicitly connects speakers’ contributions to each other (e.g. “it sounds like you are agreeing/disagreeing with Jasmine”):

2  0

☐ present  ☐ not present

Section E: Lesson Part IIIB: Partner work with novel

E1. Students work in partners with the novel and workbook pages

2  0

☐ present  ☐ not present

E2. If not present

☐ other arrangement, students working independently  1
☐ other arrangement: students working in table groups  2

E3. Students are reading the novel and/or recording responses in the workbook

2  1  0

☐ all or nearly all students  ☐ most students  ☐ a few students

E4. Students discuss passage or comprehension question for the novel with their partner or table group

2  1  0

☐ present  ☐ partly present  ☐ not present

E5. Teacher needs to intervene to manage partner work

0  1  2

☐ continually  ☐ occasionally  ☐ never or only one time
Section F: Lesson Part IV: Assign Homework

F1. Teacher assigns homework
   2  0
   □ present    □ not present

F2. Check with coach about why not: _______________________________________________

F3. Teacher uses the program slides to introduce homework
   2  0
   □ present    □ not present

F4. Teacher explicitly explains homework instructions
   2  0
   □ present    □ not present

F5. Students have their workbooks or other materials open to the relevant pages as teacher
    introduces/explains homework
   2  1  0
   □ all or nearly all students   □ some students   □ no one

Section G: Class Dynamics

G1. Across all lesson segments observed, frequency of interruptions of lesson by student behavior:
   0  1  2  3
   □ continually      □ frequently      □ intermittently      □ rarely

G2. Disruptive behavior involved:
   0  1  2  3
   □ most students    □ around half    □ around a quarter    □ a few or no students

G3. In the following space, please record qualitative observations. Include:
   • Any major teacher deviations from the lesson plan
• Any severe disruptions to the lesson (e.g. fire drill from 9:00-10:00)
• Any environmental concerns (e.g. very high noise level during fluency work)
• Any particular areas of challenge for students (e.g. some pairs did not understand the directions for fluency work)
• Any disruptions in observation (e.g. observer left room from 9:15-9:20)
<table>
<thead>
<tr>
<th>Adherence Scale</th>
<th>Quality</th>
<th>Participant responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fluency</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a=0.78 Fluency Work Happened/Did not happen</td>
<td>a=0.73 Teacher circulates during fluency work and offers support with the process</td>
<td>Students ask each other and answer follow-up questions or comment to partner</td>
</tr>
<tr>
<td>Students are grouped into partners for fluency work</td>
<td>Students are seated so that they face each other and the teacher</td>
<td>Students have materials to record new words/mark quotes</td>
</tr>
<tr>
<td>Both partners have a chance to read a passage aloud during fluency work</td>
<td>Teacher leads a summary discussion of the preceding day's guided reading passage</td>
<td>Students participate in summarizing the previous day's guided reading passage</td>
</tr>
<tr>
<td>Students record elapsed time and WPM during fluency work</td>
<td>Teacher uses a whiteboard or projector to introduce new words before reading</td>
<td>Students participate in discussing the new words for the guided reading passage they will read next</td>
</tr>
<tr>
<td>Students in the class are working in more than one fluency level, A-D</td>
<td>Teacher sets a purpose for reading the next section of the novel</td>
<td>Students provide extended responses during discussion of the novel</td>
</tr>
<tr>
<td>Students read phrase-cued passage or challenging phrases and words out loud</td>
<td>When directing students to silently read a chunk of the novel, teacher provides a context or a purpose for reading that chunk of text</td>
<td>In discussion of passage meaning, students reference text explicitly</td>
</tr>
<tr>
<td>Students record answers to comprehension questions about fluency passage</td>
<td>Teacher asks students to re-read or refer back to text</td>
<td></td>
</tr>
<tr>
<td><strong>Comprehension</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a=0.78 Students sit in a group with the teacher with copies of the guided reading book</td>
<td>Teacher asks follow-up questions to elicit fuller or clearer student responses</td>
<td></td>
</tr>
<tr>
<td>Teacher talks about the new words in a meaningful context</td>
<td>Teacher explicitly connects speakers' contributions to each other</td>
<td></td>
</tr>
<tr>
<td>Students read silently as directed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students participate in discussion of guided reading novel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher directs students to silently read particular text chunks and then stop for discussion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher poses literal (&quot;right there&quot;) questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher poses &quot;search and think&quot; questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students work in partners with the novel and workbook pages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students are reading the novel and/or recording responses in the workbook</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students discuss passage or comprehension question for the novel with their partner/table group</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B – Instructional materials

Character trait handout

<table>
<thead>
<tr>
<th>A List of Character Traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>active</td>
</tr>
<tr>
<td>adventurous</td>
</tr>
<tr>
<td>affectionate</td>
</tr>
<tr>
<td>afraid</td>
</tr>
<tr>
<td>ambitious</td>
</tr>
<tr>
<td>angry</td>
</tr>
<tr>
<td>anxious</td>
</tr>
<tr>
<td>argumentative</td>
</tr>
<tr>
<td>astounded</td>
</tr>
<tr>
<td>attentive</td>
</tr>
<tr>
<td>babyish</td>
</tr>
<tr>
<td>bewildered</td>
</tr>
<tr>
<td>bored</td>
</tr>
<tr>
<td>bossy</td>
</tr>
<tr>
<td>brown</td>
</tr>
<tr>
<td>brilliant</td>
</tr>
<tr>
<td>busy</td>
</tr>
<tr>
<td>calm</td>
</tr>
<tr>
<td>carefree</td>
</tr>
<tr>
<td>cautious</td>
</tr>
<tr>
<td>charismatic</td>
</tr>
<tr>
<td>charming</td>
</tr>
<tr>
<td>cheerful</td>
</tr>
<tr>
<td>childish</td>
</tr>
<tr>
<td>clever</td>
</tr>
<tr>
<td>clumsy</td>
</tr>
</tbody>
</table>

| | excited |
|-----------------------------|
| foar | faithfuI |
| faithful | fighty |
| fierce | foolish |
| funny | frustrated |
| friendly | frusruated |
| generous | gentle |
| gloomy | greedy |
| gloomy | grouchy |
| happy | hateful |
| helpful | helpless |
| hopeful | humourous |
| ignorant | impulsive |
| imaginative | impatient |
| impulsive | impatient |
| inquisitive | inactive |
| independent | independent |

| | puzzled |
|-----------------------------|
| quick | reliable |
| quiet | respected |
| relied | responsible |
| rested | thankful |
| rowdy | thoughtful |
| rude | thoughtless |
| sarcastic | thrifty |
| satisfied | touchy |
| scored | treasuring |
| secretive | trustworthy |
| selfish | sensitive |
| self-reliant | uncontrolled |
| sensitive | unfriendly |
| silly | unselfish |
| sincere | upset |
| skilful | useful |
| smart | valiant |
| sneaky | versatile |
| sociable | vivacious |
| snobbish | vulgar |
| sociable | warm-hearted |
| stingy | week |
| strong | wise |
| strict | wittty |
| stubborn | warm-hearted |
| studious | worried |
Narrative arc chart
Phrases for Locomotion “block party”

Day 13 Activities

1. Review homework

Students share their memory poems. Consider having partners share poems, and then asking for volunteers to read their poems to the whole class. Encourage other students to respond with something they liked about the poem.

2. Block party to launch Locomotion

Each student has a phrase strip showing a line from Locomotion. Students mingle, reading their phrases to each other and discussing how they might be related. Then, small groups work together to make predictions about the characters, setting, or happenings in the new book based on the phrases.

1. Introduce the activity.
2. Distribute one phrase strip to each student (repeats are fine).
3. Students mingle, reading their phrases to each other and discussing possible connections.
4. Class breaks into small groups of four to five students.
5. Groups discuss what they think will happen in the new book based on all the phrases they heard and ideas they discussed.
6. Groups work together to write predictions on workbook p. 79.

Locomotion phrases
- When you don’t know where your sister is anymore
- I used to laugh real loud
- Children should be seen and not heard, Miss Edna said.
- It’s Miss Edna’s over and over and over
- Be quiet!
- There used to be four of us
- I see the firemen wearing oxygen masks
- Sometimes people lie
- I don’t know why he’s so evil some days
- Stop thinking about moving ‘cause this is home
- Not a whole lot of people be saying Good, Lonnie to me
- Feels like June’s a long, long way away
- Me and her don’t always get along but she’s all I got right now
- I write the word HITS on my hand
- Knows some things we’ll never know
Student poems from creative writing activity

Kevin’s poem:
Sunny hot day
tons of people
running to the tallest ride
waiting in line like waiting for a
microwave to beep
strapped in
going uphill slowly
at the top
everyone screaming
going downhill
going upside down and all around
having fun
finally stopped

Sierra’s poem:
In the car
At a rest area
in Connecticut on my birthday
getting gas
thinking I’m going to be [inaudible],
use the bathroom,
back in the car
put on my seatbelt
my parents tell me we’re going to
Florida.
I freak out, then realize
we have 17 more hours to go,
I look back,
and that’s why
my little sister is wearing shorts
my mom told me
she packed me clothes,
we drive and drive,
I could feel it getting hotter
I have not been to Florida in years
I tell them this was the best birthday gift ever.
“Line Break Poem” activity

Ms. Marcus says line breaks help us figure out what matters to the poet. 

Don’t jumble your ideas

Ms. Marcus says Every line should count.

(Woodson, 2003, p. 4)
References


Institute of Education Sciences. (2016). Funding opportunities | search funded research grants and contracts.


Institute of Education Sciences National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.