Published in final edited form as:
*Except Child*. 2012 ; 78(3): 318–334.

# Effects of a Tier 2 Supplemental Reading Intervention for At-Risk Fourth Grade Students

**Kristen D. Ritchey[Associate Professor]**,
University of Delaware, School of Education, Willard Hall Education Building, Newark DE 19716, phone: 302-831-2325, fax: 302-831-4110 kritchey@udel.edu

**Rebecca D. Silverman[Associate Professor]**,
University of Maryland College Park, Department of Special Education, 1308 Benjamin Building, College Park, MD 20742, phone: 301-405-6465, fax: 301-314-9158 rdsilver@umd.edu

**Elizabeth A. Montanaro[Doctoral Candidate]**,
University of Maryland College Park, Department of Special Education, 1308 Benjamin Building, College Park, MD 20742, phone: 301-405-6515, fax: 301-314-9158 bethmont@umd.edu

**Deborah L. Speece[Professor]**, and
University of Maryland College Park, Department of Special Education, 1308 Benjamin Building, College Park, MD 20742, phone: 301-405-6482, fax: 301- 314-9158 dlspeece@umd.edu

**Christopher Schatschneider[Professor]**
Florida State University, Department of Psychology and Florida Center for Reading Research, Department of Psychology, 1107 W. Call Street, Office C234V, Tallahassee, Fl 32306-4301, phone: 850-644-4323, fax: 850-645-8947, schatschneider@psy.fsu.edu

## Abstract

This study investigated a Tier 2 intervention in the context of a Response to Intervention (RTI) model for 123 fourth grade students who were identified as having a high probability of reading failure. A randomized control trial was used to evaluate the effects of a 24 session multi-component supplemental intervention targeting fluency and expository comprehension of science texts. Intervention students performed significantly higher on comprehension strategy knowledge and use and science knowledge, but not on word reading, fluency, or other measures of reading comprehension. Moderators of intervention effects were also examined; children at higher risk in the intervention condition appeared to benefit more in comparison to lower probability children in intervention and compared to higher probability children in the control condition.

### Keywords

Reading Intervention; Reading Difficulties Grade 4; Response to Intervention; Tier 2 Intervention; Expository Text Comprehension

For too many children, reading failure is a persistent problem. The number of fourth-grade students reading below grade level remains unacceptably high. The National Assessment of Educational Progress reported that 34% of fourth-grade students in the United States perform below basic levels in reading (NCES, 2007). Although efforts in early identification and intervention have been successful in addressing the needs of younger students, little

Corresponding Author Contact Information Kristen D. Ritchey, Ph.D. School of Education Willard Hall Education Building University of Delaware Newark DE 19716 kritchey@udel.edu 302-831-2325.

attention has been devoted to children in later elementary grades with late emerging reading problems or who have persisting reading problems (Catts, Hogan, & Adolf, 2005; Compton, Fuchs, Fuchs, Elleman & Gilber, 2008; Leach, Scarborough & Rescorla, 2003). For many children, this means that either they encounter new problems in reading as they get older or continue to exhibit reading problems that were apparent earlier in their development.

## Response to Intervention

Effective and appropriate interventions are necessary to meet the needs of students in upper elementary grades. Recent efforts to develop alternative identification and intervention models for students with learning disabilities (LD) have led to models such as Response to Intervention (RTI). In RTI models, universal screening is used to identify students who may be at risk. These students receive tiered supplemental instruction that increases in intensity as needed. In most models of RTI, general education classroom instruction is considered Tier 1 and students who are identified as at risk receive Tier 2 instruction, which includes supplemental intervention in small groups of children several days per week. Students who are “nonresponsive” to Tier 2 instruction may continue in Tier 2 instruction or receive more intensive, individualized reading instruction (i.e., Tier 3).

The extant literature on RTI focuses on the early grades (e.g., O'Connor, Harty, & Fulmer, 2009; Vaughn, Linan Thompson, & Hickman, 2003). These studies of students at the early elementary level indicate the potential for multiple component interventions for Tier 2 and some general guidelines about the intensity of instruction (frequency and duration) that might be appropriate. However, in these studies the focus is on word level and fluency skills with little vocabulary or comprehension instruction. Older students often require attention to multiple reading domains. As such, for students in upper elementary grades and above, little is known about the optimal instructional content for a Tier 2 reading intervention.

There are few studies of multiple component interventions for students in fourth and fifth grades. In a recent review of literature of reading interventions, Wanzek, Wexler, Vaughn, and Ciullo (2010) identified only two intervention studies targeting more than one reading domain. Therrien, Wickstrom, and Jones (2006) focused their experimental study on fluency and comprehension and reported significant differences favoring the treatment group on fluency but not on a broad measure of reading that included word identification and passage comprehension. The intervention was delivered over 50 sessions of 10 to 15 minutes each (for 8 to 12 hours of instruction). O'Connor, Bell, Harty, Larkin, Sackor, and Zigmond (2002) provided approximately 36 hours of instruction to two experimental groups who received instruction in phonological awareness, phonics, reading text, fluency, and comprehension. The authors manipulated the reading level of text used for instruction. Both treatment groups outperformed the control group on second grade level fluency, word reading, and comprehension measures; the groups did not differ on listening comprehension or fluency in fourth grade level text.

Adoption of a validated program within an RTI framework is challenging for this age range, given the paucity of studies for upper elementary grade students. The two reviewed studies (O'Connor et al., 2002; Therrien et al., 2006) used narrative text. As children advance through the elementary grades they are increasingly confronted with the need to access and comprehend expository text. The complexity of expository text may pose a particular challenge for children who have reading difficulties. We did not identify any experimental studies that used a multi-component intervention with a focus on expository text for upper elementary struggling readers. In the following section, we describe components of effective remedial intervention for students in upper elementary school that are potential targets for supplemental intervention.

## Components of Tier 2 Interventions for Upper Elementary Grade Readers

Comprehension is the ultimate goal of reading instruction. The focus of remediation in older readers can be complex, since poor comprehension can be caused by deficits in decoding, vocabulary, fluency, motivation, or background knowledge, as well as a lack of strategic behaviors for monitoring and repairing misunderstanding (Snow, 2002). Since readers may struggle in more than one area, a multi-component intervention may be most appropriate within Tier 2. Results from previous studies provide evidence that multi-component interventions may be effective for older struggling readers (O'Connor et al., 2002; Therrien et al., 2006). These findings are supported by a recent meta-analysis of interventions for older readers indicating that multi-component interventions can result in significant gains (Edmonds et al., 2009).

Although some upper elementary grade children have not mastered beginning decoding skills, a much larger group of students lacks the decoding strategies necessary to attack multisyllabic words (Archer, Gleason, & Vachon, 2003). Instruction in decoding multisyllabic words can have positive outcomes for struggling readers in fourth through eighth grades, including improved word recognition and comprehension (Diliberto, Beattie, Flowers, & Algozzine, 2009). Breaking multisyllabic words into smaller parts can also assist readers in determining the meaning of unfamiliar words, including using prefixes, suffixes, and root words to determine meaning (Cunningham, 1998). Vocabulary knowledge, including morphological awareness, is related to reading comprehension (Beck, Perfetti, & McKeown, 1982; Nagy, Berninger, & Abbott, 2006), as comprehension depends on understanding the meanings of between 90 and 95% of the words in a text (Nagy & Scott, 2000). Beck, McKeown, and Kucan (2002) suggest teaching students useful words that frequently appear across content areas, presenting them in context, providing “student friendly” definitions, and repeatedly exposing students to words.

Students also need to be able to read fluently. Fluent oral reading has been shown to have a strong positive relationship with reading comprehension (Pinnell, Pikulski, Wixson, Campbell, Gough, & Beatty, 1995). Repeated reading is a commonly used intervention for improving reading fluency (Chard, Vaughn & Tyler, 2002). Significant gains in fluency and comprehension have been found using repeated reading interventions (O'Connor, White, & Swanson, 2007; Therrien & Hughes, 2008).

Students in upper elementary grades are also faced with the challenge of the shift from reading primarily narrative to reading primarily expository texts (Grigg, Daane, Jin, & Campbell, 2003). In particular, struggling readers have difficulty identifying the most important information presented in a text (Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007). These students require explicit instruction and modeling in the use of multiple comprehension strategies (Gersten, Fuchs, Williams & Baker, 2001). Effective interventions for improving expository comprehension have included teaching text structures (Englert & Mariage, 1991), self-monitoring (Graves, 1986), and understanding Question Answer Relationships (QAR; Raphael & Au, 2005).

In designing an intervention for readers who may have been struggling for several years, it is essential to include components aimed at increasing motivation. Students who have experienced repeated failure are likely to have less motivation to read or put effort into learning new strategies for reading (Minskoff, 2005; Morgan, Fuchs, Compton, Corday, & Fuchs, 2008). Guthrie and his colleagues (2004a, 2004b) used an approach to reading instruction called Concept-Oriented Reading Instruction, which increased both reading comprehension and motivation in typically developing children. Motivational components of this approach include (a) embedding reading instruction within a content area, such as

science; (b) selecting interesting texts; (c) providing opportunities for students to make choices; and (d) including hands-on activities related to the texts.

In sum, designing an intervention for students in upper elementary grades is complex. It must provide strategies for reading multisyllabic words, include vocabulary instruction in context, and reading fluency practice. These can be combined with instruction on specific strategies for comprehending expository texts, while including components in the intervention to attend to the motivational needs of students who may have experienced repeated reading failure.

## Assessing Responsiveness for Intermediate Grade Students

In the context of RTI, defining responsiveness is key. As typically conceptualized, Tier 2 interventions are designed to be relatively brief (e.g., 8 to 12 weeks) to determine who may need a more intensive intervention (i.e., Tier 3) and who received enough of a boost as to no longer require intervention beyond Tier 1. Responsiveness to instruction can be defined by level of achievement attained, by amount of growth over time, or both (Fuchs & Fuchs, 1998). Curriculum-based measures were incorporated into the present study to assess growth as the intervention included some attention to fluency. Near transfer and far transfer measures also require consideration. Near transfer measures that reflect instructional emphases may be best suited for short-term interventions, but far transfer measures like published, nationally-normed tests may provide more generalizable results. We used both types of measures in this study.

## Context of the Current Study and Research Questions

The current study is based on work from the second and third years of a five-year longitudinal project studying reading disabilities in middle childhood, specifically fourth grade. In the first year, the screening procedure used to identify at-risk fourth grade students was validated (Speece, Ritchey, Silverman, Schatschneider, Walker, & Andrusik, 2010). We applied the screening procedure to identify fourth grade students who had a higher probability of reading failure in two consecutive cohorts of students. The purpose of this study is to examine the effects of a supplemental, multi-component Tier 2 intervention for fourth grade students who have a high probability of reading failure. We developed an intervention using science content texts, positing that such an intervention could be motivating for students given interesting texts and could ease scheduling demands of supplemental intervention that often require students to miss content area instruction. The research questions investigated are: (1) What are the effects of a supplemental reading intervention for fourth grade students identified as having a higher probability of reading failure compared to children receiving typical classroom instruction? (2) Does initial predicted probability of reading failure and other moderators of reading predict responsiveness to a supplemental reading intervention? (3) Does initial predicted probability of reading failure and other moderators reduce the probability of reading failure after intervention?

## Method

### Participants

The participants were 123 fourth grade students identified as having a higher probability for reading failure compared to their classmates (selection procedures described below). The mean age of students was 9 years, 7 months ($SD$ = 4.93 months). No students received special education services for any academic area. Demographics for the screening sample, intervention group, and control group are found in Table 1. There were no significant

differences between the intervention and control groups for gender, $\chi^2$ (1, $N$ = 123) = .004, $p$ = .95, ethnicity (collapsed into Black, White, and Other categories) $\chi^2$ (2, $N$ = 120) = 2.00, $p$ = .37, or mother's education, $\chi^2$ (3, $N$ = 119) = .32, $p$ = .96. Two students withdrew from participation during the intervention; attrition analysis was not possible given the sample cell size. There were no differences by cohort.

**Instructional Setting—**The students were enrolled in one of 11 parochial schools in the Mid Atlantic region. There were 18 unique teachers in the study, and eight teachers were in the study both years. On average, teachers had 15.89 years of teaching experience ($SD$ = 11.72, range 1 to 40 years) and 7.44 years teaching fourth grade ($SD$ = 7.75; range of 1 to 26 years). Teachers identified their race as Black ($n$ = 3), Multiracial ($n$ = 1), and White ($n$ = 14). One teacher had a bachelor's degree; eleven had a bachelor's degree plus some credits toward a master's degree; five had a master's degree; and one had a master's plus 30 credits. Nine teachers had certification, and five were pursuing state certification.

The parochial schools were part of a large group of schools coordinated by a central administration. Individual schools selected a core reading curriculum and made site-based decisions about instruction. To document the quality of Tier 1 instruction classroom teachers were observed during reading and language arts twice during the school year. Observers took field notes, and at the end of the observation session, rated teachers on a scale of 0 (low quality) to 3 (high quality) on three facets of instruction: delivery, management, and content. Ratings from the indicators were averaged to obtain an overall instructional quality rating. High quality ratings were given when instruction aligned with research-based practices. Inter-rater agreement for observers was above .90. The mean ratings were 2.64 ($SD$ = .21) across the two years (min = 2.11; max = 2.90). These data suggest Tier 1 instruction was of satisfactory instructional quality.

## Measures

The following section describes the measures. There were five pre-test and post test measures, two progress monitoring measures, three post-test only measures, and two moderator variables.

**Word Recognition and Decoding—**The Woodcock Johnson Tests of Achievement, Third Edition (WJ III) (Woodcock, McGrew, Mather, & Schrank, 2001) Letter Word Identification and Word Attack subtests were administered. The Letter Word Identification subtest assesses word recognition skills. The mean split-half reliability coefficient is .94. The Word Attack subtest assesses students' ability to decode phonetically-regular pseudowords as an evaluation of phonetic and structural analysis skills. The split-half reliability coefficient is .87.

**Gates-MacGinitie Reading Test, Fourth Edition (GMRT)—**The GMRT Reading Comprehension subtest (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) was administered. Students are given 35 minutes to silently read short narrative and expository passages and answer multiple choice questions. The alternate form and the test-retest reliability exceed . 90 for fourth grade students.

**Test of Word Reading Efficiency (TOWRE)—**The Sight Word Efficiency (SWE) and Phonemic Decoding Fluency (PDF) subtests of the TOWRE (Torgesen, Wagner, & Rashotte, 1999) were administered to assess real word and nonword reading efficiency skills. SWE assesses fluent reading of real words, and PDE subtest assesses fluent decoding of nonsense words. The authors report excellent alternate-form reliability ($r$ =.93) and strong concurrent criterion-related validity ($r$ = .87 to .89).

**Maze—**Maze (Fuchs, n.d.) was administered as an assessment of silent reading and comprehension. The task uses a modified cloze technique. The first sentence of a reading passage remains intact, and every seventh word thereafter is deleted and replaced with three choices. The student selects the choice that is appropriate in the context. Students are given 2 min to complete as many choices as possible. The mean number of correct choices for two probes was converted to items correct per minute. The reliability and criterion validity of the Maze are strong ($r$ = .60 to .86; Fuchs & Fuchs, 1992).

**Passage Reading Fluency (PRF)—**PRF (Fuchs, Hamlett, & Fuchs, 1990) was administered as an assessment of oral reading fluency. Students are given one minute to read a narrative passage. Two fourth grade passages were administered and the mean words correct per minute was calculated. Reliability (test-retest, alternate form) exceed .90 across studies, and criterion validity is strong (Deno, 1985; Fuchs & Fuchs, 1992).

**Word Identification Fluency (WIF)—**WIF was developed for this project and was administered as an assessment of word reading fluency. WIF was developed based on the procedure used by D. L. Compton (personal communication, March 3, 2003). Words on WIF were randomly selected from the Educator's Word Frequency Guide (Zeno, Ivens, Millard, & Duvvuri, 1995) that represented words with a range of frequency levels. Parallel probes of 80 words each were developed. The mean number of correct choices for two probes was converted to items correct per minute. In the screening validation sample (Speece et al., 2010), the parallel forms reliability coefficient was .92. Validity coefficients with the WJ III Word Identification subtest ($r$ = .68), TOWRE SWE ($r$ = .86), and PRF ($r$ = .78) are strong.

**Assessment of Strategy Knowledge and Use for Information Text (ASKIT)—**To assess students' knowledge of and ability to use comprehension strategies for information text, the Assessment of Strategy Knowledge and Use for Information Text (ASKIT, Form A) was administered. In this researcher-developed assessment, students were asked to answer questions about reading strategies (previewing, identifying the main idea, retelling, summarizing) and to demonstrate these reading strategies through reading an authentic information text (O'Sullivan, 2003). Students read the text orally and miscues were counted. Each knowledge and strategy question was scored using a 0 to 3 scale, and criteria and anchor points for each score were defined. Two raters scored all items, inter-rater agreement exceeded .90, and any discrepancies were discussed to yield a final score. Two scores were derived: percentage of words read correctly (untimed Reading Accuracy) and total raw score (Comprehension). Cronbach's alpha for this sample was .56 ($n$ = 123), just below the recommended .60 for researcher-developed measures (Gersten, Fuchs, Compton, Coyne, Greenwood, & Innocenti, 2005). We judge this to be acceptable for research purposes given a likely restricted range with this at-risk sample.

**Science Knowledge—**To assess learning of science content that was part of the intervention, a 28 item multiple-choice test was administered. The test items assessed knowledge of science concepts (e.g., physical and behavioral adaptations, interdependence) and characteristics of forest animals (e.g., habitats, predators/prey, classification of animals) that were included in the intervention texts. Test items were read to students. The internal consistency of the Science Knowledge assessment was .71 ($n$ = 123).

**Academic Competence—**The Academic Competence subscale of the Social Skills Rating System (SSRS; Gresham & Elliott, 1990) was completed by classroom teachers. The subscale has nine items which require the rater to use a three-point scale to compare the child to classmates on reading, math achievement, and motivation to learn. The authors

reported excellent test-retest reliability for the teacher form of the Academic Competence subtest ($r$ = .93). The criterion validity is moderate to strong with other teacher rating measures.

**Extra services—**The number of extra services provided to students was collected to account for other academic support that students may have received (Case, Speece, & Molloy, 2003). These included small group interventions in reading or math, special education services, Title 1 services, speech and language therapy, counseling, or out-of-school tutoring. The total number of extra services was used in analyses (max = 10). We interpret scores on this measure to reflect the degree of concern school personnel have for a student rather than as an indication of the impact of additional services on achievement.

## Procedure

**Screening, selection and assignment to condition—**After research activities were approved by the Institutional Review Board for Research with Human Subjects, permission letters that included an informed consent form were sent to all enrolled fourth grade students in 11 schools. In the fall, 463 students (251 students in cohort one, 212 students in cohort two) were screened. Students were selected for intervention based on screening procedures developed in the first year of the project (Speece et al., 2010). The screening battery included GMRT Reading Comprehension, Test of Silent Word Reading Fluency (Mather, Hammill, Allen, & Roberts, 2004) and Teacher Reading Rating (TRR, Speece et al.). For the TRR, students are rated on a 1 to 5 scale, and for students rated as 1 or 2 (below grade level), teachers identified the number of problem areas experienced by the student (decoding/word reading, fluency, comprehension, vocabulary, motivation). Raw scores from GMRT Reading Comprehension, TOSWRF, and the number of reading problems from the TRR were entered into a logistic regression equation to determine the probability of reading risk. Students with a predicted probability of risk ≥ .40 were identified as the initial pool of participants. A higher probability indicates higher risk. The decision to select .40 as the criterion was a practical one in that it provided the number of students for whom we had the resources to include in the assessment and intervention activities.

Students were rank ordered by predicted probability within school. Four or more students at each school were needed to have sufficient students to form intervention and control groups. Students were paired by initial predicted probability and randomly assigned to the intervention condition ($n$ = 57) or a nonintervention control condition ($n$ = 66). After intervention groups were formed, initial progressing monitoring assessments were administered in January of each year. Four progress monitoring time points were selected to estimate growth at approximately three week intervals; four points allowed for estimation of nonlinear growth.

**Intervention—**Intervention was implemented for two consecutive years using the same standard protocol. The intervention consisted of 24 scripted lessons implemented over 12 to 15 weeks (mid-January to April). Intervention was provided three 40-minute sessions per week 16 hours total) in groups of two to four students. Intervention was provided in addition to general reading instruction provided by the classroom teachers, and scheduled at the teachers' preferred time to the extent possible. Any additional services that students were getting continued. Graduate research assistants, some of whom had previous teaching experience, served as tutors. There were 12 tutors total; one tutor was an instructor for both years. Tutors participated in approximately 20 hours of training and demonstrated fidelity prior to intervention.

The focus of the intervention was expository text comprehension, and science texts were used. Two units of instruction (animals of temperate forests and tropical rain forests) were developed and identified using state curriculum standards for reading and life sciences. Lesson plans, the list of text used, and the scope and sequence are available from the first author.

**Fluency:** Students engaged in repeated reading using a passage read in the previous lesson, and was 5-7 minutes of the session. First, the tutor modeled fluent reading of the text selection (Chard et al., 2002). Next, students engaged in repeated reading individually or with a partner. Each lesson alternated between students rereading the passage for 3 minutes individually and rereading the passage with a partner (2 minutes per student). Students read the passage as many times as possible within the total time period. When reading with a partner, students provided peer feedback on miscues consistent with Peer Assisted Learning Strategies (PALS) procedures (Fuchs, Fuchs, Mathes, & Simmons, 1997). Additionally, the tutor listened to one student during each session and provided feedback using a standardized feedback protocol that included telling the student the number of words read correctly, providing a compliment, and making a recommendation for improvement on a specific skill.

**Comprehension:** Each lesson included explicit comprehension instruction, vocabulary instruction, and text instruction was approximately 25-30 minutes per session. Five strategies were targeted for instruction: (a) previewing expository texts, (b) monitoring for understanding, (c) using strategies for decoding unfamiliar words, (d) finding the main idea using paragraph shrinking (Fuchs et al., 1997; Jenkins, Heliotis, Stein, & Haynes, 1987), and (e) question and answer relationships (QAR) (Raphael & Au, 2005). When teaching these strategies, the tutor provided cognitive modeling and multiple opportunities for the students to practice the strategy, both in isolated text examples and within authentic texts.

**Vocabulary:** Two to four words were introduced in each lesson. Tutors followed the instructional protocol for teaching vocabulary in context, as described in Beck et al. (2002). Tutors referenced how the word was used in the text by repeating or paraphrasing the sentence in which the word first appeared. Then, they provided students with a definition, and the tutor and students generated examples and non-examples, and extended use of the word to other contexts.

**Text instruction:** Instructional time was allocated to reading and discussing authentic texts. Texts were identified by reviewing commercially-available texts and identifying text with an appropriate instructional reading level (late second grade to third grade). Texts were also selected for interest and age appropriateness. Tutors modeled fluent reading, and students read the text orally and silently. The lesson plans directed tutors to model comprehension strategies, introduce new vocabulary, and ask students text-based questions. Students also practiced applying comprehension strategies within connected text.

**Motivational components:** Motivational components were included in the intervention (Guthrie et al., 2004a, 2004b). Students were provided with opportunities to make choices. During some sessions of repeated reading, students could select one of two passages or students were given opportunities to select and read a preferred chapter or section of text. Additionally, four hands-on science activities were interspersed throughout the intervention, and were a part of several intervention sessions.

**Fidelity of intervention implementation—**Fidelity of intervention was monitored in several ways. Fidelity criteria were identified for each component of the lesson (points per lesson ranged from 25 to 52) and were included on the lesson plan to guide tutors in

implementing the lesson plan. Prior to providing instruction, tutors demonstrated that they could implement randomly-selected lessons with at least 90% of the components. Tutors were also observed while implementing the intervention by project staff, and corrective feedback was provided as needed.

All lessons were audio-recorded, and approximately 25% of lessons ($n$ = 123) across tutors and sessions were evaluated for fidelity. The mean time of implementation for the evaluated lessons was 40 min, 40 sec ($SD$ = 1 min 47 sec). Approximately 93% ($SD$ = 4%) of the components were implemented, and there were no significant differences in total fidelity by year, $F(121) = .05$, $p = .823$, or implementation time by year, $F(120) = .252$, $p = .617$.

**Other intervention variables—**Attendance data were collected to determine the number of intervention sessions attended. Tutors rated each student's attention during the intervention session on a 1 to 5 scale (1 = very poor to 5 = very good, required no prompting).

**Data analyses—**Differences between the control and intervention conditions (Group) were investigated through multilevel modeling with students nested within classroom (Littell, Milliken, Stroup, & Wolfinger, 1996). Pre-post test differences, post-test only differences, and growth differences were analyzed. Classroom-level intercepts were allowed to vary and were included as random effects (22 of 24 instructional groups were formed within classroom). Intraclass correlations for classrooms ranged from 0 to .186. Group was included as a fixed effect. For pre-test/post-test variables, differences between groups were estimated with two-level modeling with pre-test as a control variable to increase precision of estimates and to account for any differences between groups that may have been present [i.e., random effects analysis of covariance (ANCOVA)]. For post-test only measures, differences between groups were estimated via two-level hierarchical linear models with post-test as the outcome [i.e., random effects analysis of variance (ANOVA)]. Benjamani-Hochberg corrections for multiple comparisons were made.

For progress monitoring variables, differences between the intervention and control group were estimated via three-level linear growth models with time nested in students nested in classroom (Singer, 1998). The intercept for progress monitoring analyses was centered at the last measurement occasion to estimate group differences at the end of intervention and rate of growth across time. The first time point was used as a covariate in progress monitoring analyses to control for initial group pre-test differences. Covariates were grand-mean centered. The effect size was calculated using Hedges' $g$. The effects of potential moderators of learning were examined by estimating interactions between these moderators and Group. The moderator variables included extra services, Academic Competence, and initial predicted probability of reading failure. Exploratory analyses investigated intervention-only predictors including tutor ratings of students' attention, intervention attendance rate, and tutor fidelity to the intervention protocol.

## Results

Table 2 includes descriptive statistics for pre-test and post-test measures, post-test only measures, progress monitoring measures across four time points, and moderators. Standard scores for norm-referenced measures are reported for descriptive purposes, but raw scores were used in analyses. Table 2 also includes descriptive statistics for initial predicted probability (i.e., probability of having a reading problem at the beginning of fourth grade) and posterior predicted probability (i.e., probability of having a reading problem at the end of fourth grade).

### Differences on Pre-Post Measures

Table 3 presents the Group × Time results for the random effects ANCOVA models on the following measures: TOWRE Sight Word Efficiency, TOWRE Phonemic Decoding Efficiency, WJIII Letter Word Identification, WJIII Word Attack, GMRT Reading Comprehension, and Maze. There were no significant effects for Group × Time. We also investigated moderators including Academic Competence, extra services, and initial predicted probability. We found no significant effects in these analyses.

### Differences on Post-Test Only Measures

There were three post-test only measures: ASKIT Reading Accuracy, ASKIT Comprehension, and Science Knowledge (see Table 3). There was a significant effect of Group on ASKIT Comprehension, $F(1, 16) = 10.09$, $p = .006$, favoring the intervention group, $g = .5631$. This was statistically significant after correcting for multiple comparison There was a significant effect of Group on Science Knowledge, $F(1, 16) = 12.70$, $p < .0031$, favoring the intervention group, $g = .6458$.

The effect of intervention on probability of reading failure was determined. The posterior predicted probability for the control group was .69 ($SD = 0.05$) and for the intervention group was .57 ($SD = 0.06$). There was no significant difference in posterior predicted probability by Group, controlling for initial predicted probability, $F(1, 102) = .74$, $p = .40$.

We investigated moderators in these analyses to determine whether there were interactions between Group and Academic Competence, extra services, and initial predicted probability. There was a Group × extra services interaction, $F(1, 101) = 5.49$, $p < .05$, on ASKIT Comprehension (see Table 3). Post hoc analysis showed no effect of the intervention over control, $g = 0.17$, for students with fewer extra services (one $SD$ below the mean), but a substantial effect, $g = 1.01$, for children with a higher number of extra services (one $SD$ above the mean). Children who received more services and who received the intervention outperformed the students in the control group. There were no other significant effects.

### Differences on Growth Measures

Unconditional models were estimated to determine whether there was significant variance in intercept and slope for progress monitoring measures (PRF and WIF). The model for PRF produced a significant random intercept and slope. WIF had significant intercept and significant quadratic parameter. Conditional models were constructed to test differences in growth by group (see Table 4). A significant interaction was found for WIF, $F(1, 584) = 3.97$, $p < .05$. There was no difference between the mean of the control and the intervention groups at any of the four time points. At the last data point, the control group's score decreases slightly and the intervention group's scores increased slightly, which may be responsible for the interaction. No significant interaction was found for PRF.

Next, we tested the effect of moderators on conditional models using initial predicted probability, Academic Competence, and extra services. We found a significant Group × Time × initial predicted probability interaction on PRF, $F(2, 582) = 3.37$, $p < .05$ (see Table 4). Post hoc analyses of the control and intervention conditions within high and low probability showed no significant differences in the predicted scores at each time point, controlling for initial PRF and classroom teacher. It appears that differences between high and low initial predicted probabilities within the control group may be responsible for the interaction effects.

### Analysis of Intervention Moderators

We explored whether intervention variables applicable to only the intervention group (attendance, tutor rating of attention, and intervention fidelity) predicted response as defined by pretest-posttest, post test only, and growth measures. We found no significant results on these predictors, but found marginal effects of tutor ratings of attention on GMRT Reading Comprehension, $F(1, 38) = 3.75$, $p = .07$, and Maze, $F(1, 38) = 3.14$, $p = .09$, suggesting that attention may play a role in responsiveness to intervention.

## Discussion

The purpose of this study was to determine if a short-term, multi-component Tier 2 reading intervention that focused on comprehension of expository text affected the reading skills and probability of reading problems of fourth grade children within a randomized control trial. There are few studies that investigate the effectiveness of reading interventions for upper elementary children who are struggling readers, and no multi-component studies that that focus on comprehension of expository text (Wanzek et al., 2010).

The results of this experimental study are mixed. Children in the intervention group performed significantly better on the identification and application of comprehension strategies ($g = .56$) and on science knowledge ($g = .65$), both closely aligned with instruction. The effect for reading comprehension was moderated by the number of additional services received by children such that children in the intervention group who received more services benefited more from the intervention than did their counterparts in the control condition. Passage reading fluency was moderated by the number of extra services received. Together, these results suggest that children at higher risk may be the best candidates for the intervention tested here. There were no other differential effects for fluency, word level skills, broader measures of reading comprehension, or reduction of risk. There was a suggestion that the intervention was differentially effective on reading comprehension for intervention children based on the rating of attention by tutors which supports findings reported by Stage, Abbott, Jenkins and Berninger (2003). Given that fidelity of implementation was high, further research may be needed to determine the instructional content and conditions (including total amount of instruction) that may be needed to positively affect broad reading outcomes. It may be the case, as suggested by McKeown, Beck, and Blake (2009) that strategy instruction could have a lesser role as compared to content instruction.

The primary finding of this investigation, that a short-term, multi-component reading intervention is effective on a near transfer measure of comprehension, raises several issues regarding the applicability of the RTI model with upper elementary grade students. We did not identify any studies that used a multi-component approach to reading comprehension with expository text for struggling readers, but reasoned that instruction that targeted fluency, vocabulary, motivation, and comprehension skills was important for poor fourth grade readers because of evidence suggesting that children at this age likely exhibit multiple reading problems (e.g., Snow, 2002). This group of students did not demonstrate significant weaknesses in word reading skills, as evidenced by average range reading standard scores on norm-referenced assessments, also supporting the choice of intervention content. A focus on science provided an opportunity to meet scheduling conflicts which can result in students missing content area instruction in order to receive reading intervention. The effects for the science knowledge assessment indicate that students learned content during an intervention with the primary goal of improving reading.

How best to capture responsiveness is an important measurement concern within RTI models. Several of the single component studies included in the Wanzek at al. (2010) review

found significant effects for expository text comprehension. However, they did not use norm-referenced measures but instead relied on researcher-developed measures (e.g., Mason, 2004; Miranda, Villaescusa, & Viadl-Abarca, 1997). O'Connor et al. (2002) found significant effects on several nationally-normed assessments but the focus was narrative text delivered over approximately 33 hours of instruction, twice the time in the current study. Rosenshine and Meister (1994) identified the tension between researcher-developed and nationally-normed assessments in their review of reciprocal teaching and it has been a common finding subsequently (e.g., Simmons et al., 2010). It seems sensible to target measures more proximal to instruction especially when interventions are short term given that nationally-normed measures are not designed to capture growth over short time periods. It may be more appropriate to label our ASKIT measure as a "medium" transfer task given that the text used was in the same genre as instructional texts but novel to the children but the questions on knowledge and use were derived from the intervention. This type of hybrid task, when fully developed psychometrically, may serve as a bridge between near and far transfer instruments. However, in the current study we did not find uniform significant effects between intervention and control groups on fluency measures that are known to capture short-term growth (i.e., PRF, WIF). At first glance this was a puzzling result given that children spent most of their time reading text with a small but explicit emphasis on reading fluently. The measures, though, are more properly thought of as far transfer measures as the words and text were not related to science content and were narrative passages. Future research should examine fluency and accuracy measures that use words that overlap with instructional text in short-term interventions.

This approach to assessment and responsiveness may provide a middle ground for validation of short-term comprehension interventions as called for in most RTI models. Any discussion of proximal and distal measures of responsiveness should recognize there will always be a question of generalizability of findings when they do not include broader measures that researchers, teachers, and parents may prioritize. A longitudinal strategy across years may be one way to untangle these critical outcomes.

## Limitations

A primary limitation is that while we obtained information on the quality of general classroom instruction, we do not have detailed data on daily classroom practice in reading or science. To better contextualize the Tier 1 instruction that students in Tier 2 are also experiencing, more detailed data on Tier 1 instruction may be needed. Additionally, though representative of many public and private school districts around the country, the parochial schools in which the study was conducted may have had less dramatic variation in student ability levels than in other studies. This may have reduced the possibility of affecting differences between the intervention and control group. We note that our intervention sample, compared with the screening sample, had more males (59.6% vs. 49.2%), fewer Caucasian children (27.3% vs. 43.8%), and fewer mothers with at least a college education (27.3% vs. 43.8%). Also, these children were, on average, a standard deviation below national norms on reading comprehension and scored just about the 25$^{th}$ percentile on PRF winter norms provided by Hasbrouk and Tindal (2006). While we included motivational components within the intervention, we did not include a motivation outcome. Finally, because the intervention was multi-component and complex, it is impossible to know the effect of specific aspects of the intervention.

## Conclusions

Given the results, our perspective on RTI with upper elementary grade children in the area of reading comprehension is one of cautious optimism tempered by the effort required to produce what some may regard as modest effects. It may be ambitious to think that 16 hours

of instruction will unbend the twig of persisting reading failure. We did uncover some intriguing findings that showed children in intervention were able to apply their comprehension knowledge to a text not used in intervention, that this effect was moderated by the number of extra-classroom services received. However, there were other distal measures for which we did not find effects despite targeting fluency and word level skills with an overall emphasis on comprehension. Thus, while, in principle, RTI models may be applicable to upper elementary school children, in practice, more work is needed on how the goals of RTI can be achieved.

## Acknowledgments

This work was supported by grants from the National Institute of Child Health and Human Development (Grant No. P5HD052121) and the U. S. Department of Education (Grant No. H325D070082).

## References

Archer AL, Gleason MM, Vachon VL. Decoding and fluency: Foundation skills for struggling older readers. Learning Disability Quarterly. 2003; 26(2):89–101. doi:10.2307/1593592.

Beck, IL.; McKeown, MG.; Kucan, L. Bringing words to life: Robust vocabulary instruction. The Guilford Press; New York: 2001.

Beck IL, Perfetti CA, McKeown MG. Effects of long-term vocabulary instruction on lexical access and reading comprehension. Journal of Educational Psychology. 1982; 74:506–521. doi: 10.1037/0022-0663.74.4.506.

Case LP, Speece DL, Malloy DE. The validity of a Response-to-Instruction paradigm to identify reading disabilities: A longitudinal analysis of individual differences and contextual factors. School Psychology Review. 2003; 32:557–582.

Catts, HW.; Hogan, TP.; Adlof, SM. Developmental changes in reading and reading disabilities.. In: Catts, H.; Kamhi, A., editors. Connections between language and reading disabilities. Erlbaum; Mahwah, NJ: 2005.

Chard DJ, Vaughn S, Tyler BJ. A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. Journal of Learning Disabilities. 2002; 35:386–406. doi:10.1177/00222194020350050101. [PubMed: 15490537]

Compton DL, Fuchs D, Fuchs LS, Elleman AM, Gilbert JK. Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. Learning and Individual Differences. 2008; 18:329–337. doi: 10.1016/j.lindif.2008.04.003.

Cunningham PM. The multisyllabic word dilemma: Helping students build meaning, spell, and read “big” words. Reading and Writing Quarterly: Overcoming Learning Difficulties. 1998; 14:189–218. doi:10.1080/1057356980140204.

Diliberto JA, Beattie JR, Flowers CP, Algozzine RF. Effects of teaching syllable skills instruction on reading achievement in struggling middle school readers. Literacy Research and Instruction. 2009; 48:14–27. doi:10.1080/19388070802226253.

Edmonds MS, Vaughn S, Wexler J, Reutebuch C, Cable A, Tackett KK, Schnakenberg JW. A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. Review of Educational Research. 2009; 79:262–300. doi: 10.3102/0034654308325998. [PubMed: 20072704]

Englert CS, Mariage TV. Making students partners in the comprehension process: Organizing the reading “posse.”. Learning Disability Quarterly. 1991; 14:123–138. doi:10.2307/1510519.

Fuchs, LS. Project PROACT Maze Reading Passages. Vanderbilt University; Nashville, TN: n.d.

Fuchs LS, Fuchs D. Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. Learning Disabilities Research & Practice. 1998; 13:204–219.

Fuchs LS, Fuchs D. Identifying a measure for monitoring student reading progress. School Psychology Review. 1992; 21:45–58.

Fuchs D, Fuchs LS, Mathes P, Simmons D. Peer-assisted learning strategies: Making classrooms more responsive to diversity. American Educational Research Journal. 1997; 34(1):174–206. doi: 10.2307/1163346.

Fuchs, LS.; Hamlett, CL.; Fuchs, D. Monitoring basic skills progress: Basic reading. ProEd; Austin, TX: 1990.

Fuchs LS, Fuchs D, Maxwell L. The validity of informal reading comprehension measures. Remedial and Special Education. 1988; 9(2):20–29. doi:10.1177/074193258800900206.

Gersten R, Fuchs LS, Compton DL, Coyne M, Greenwood C, Innocenti MS. Quality indicators for group experimental and quasi-experimental research in special education. Exceptional Children. 2005; 71:149–164.

Gersten R, Fuchs LS, Williams JP, Baker S. Teaching reading comprehension strategies to students with learning disabilities: A review of research. Review of Educational Research. 2001; 71(2): 279–321. doi:10.3102/00346543071002279.

Graves AW. Effects of direct instruction and metacomprehension training on finding main ideas. Learning Disability Research. 1986; 1:90–100.

Gresham, FM.; Elliott, SN. The Social Skills Rating System. American Guidance Service; Circle Pines: 1990.

Grigg, WS.; Daane, MC.; Jin, Y.; Campbell, JR. The Nation's report card: Reading 2002. U.S. Department of Education, Institute of Education Sciences; Washington, DC: 2003. http://nces.ed.gov/nationsreportcard/pubs/main2002/2003521.asp

Guthrie JT, Wigfield A, Barbosa P, Perencevich KC, Taboada A, Davis MH, Scafiddi NT, Tonks S. Increasing reading comprehension and engagement through Concept-Oriented Reading Instruction. Journal of Educational Psychology. 2004; 96:403–423. doi: 10.1037/0022-0663.96.3.403.

Guthrie, JT.; Wigfield, A.; Perencevich, KC. Scaffolding for motivation and engagement in reading.. In: Guthrie, JT.; Wigfield, A.; Perencevich, KC., editors. Motivating reading comprehension: Concept-oriented reading instruction. Lawrence Erlbaum Associates Publishers; Mahwah, NJ: 2004. p. 55-86.

Hasbrouck J, Tindal G. Oral reading fluency norms: A valuable assessment tool for reading teachers. The Reading Teacher. 2006; 59:636–644.

Jenkins JR, Heliotis JD, Stein ML, Haynes MC. Improving reading comprehension by using paragraph restatements. Exceptional Children. 1987; 54:54–59. [PubMed: 3653217]

Leach J, Scarborough H, Rescorla L. Late-emerging reading disabilities. Journal of Educational Psychology. 2003; 95:211–224.

Littell, RC.; Milliken, GA.; Stroup, WW.; Wolfinger, RD. SAS system for Mixed Models. SAS Institute, Inc.; Cary, North Carolina: 1996.

MacGinitie, WH.; MacGinitie, RK.; Maria, K.; Dreyer, LG. Gates-MacGinitie Reading Tests, Fourth Edition, Forms S and T. Riverside Publishing; Itasca, IL: 2000.

Mason L. Explicit self-regulated strategy development version reciprocal questions: Effects on expository reading comprehension among struggling readers. Journal of Educational Psychology. 2004; 96:283–296.

Mather, N.; Hammill, DD.; Allen, EA.; Roberts, R. Test of Silent Word Reading Fluency: Examiner's manual. Pro-Ed; Austin, TX: 2004.

McKeown MG, Beck IL, Blake R. Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approached. Reading Research Quarterly. 2009; 44:218–253. http://dx.doi.org/10.1598/RRQ.44.3.1.

Minskoff, E. Teaching reading to struggling learners. Brookes; Baltimore, MD: 2005.

Miranda A, Villaescusa MI, Vidal-Abarca E. Is attribution retaining necessary? Use of self-regulated procedures for enhancing the reading comprehension strategies of children with learning disabilities. Journal of Learning Disabilities. 1997; 30:513–520. doi: 10.1177/002221949703000506. [PubMed: 9293233]

Morgan PL, Fuchs D, Compton DL, Cordray DS, Fuchs LS. Does early reading failure decrease children's reading motivation? Journal of Learning Disabilities. 2008; 41:387–404. doi: 10.1177/0022219408321112. [PubMed: 18768772]

National Center for Education Statistics. National Assessment of Educational Progress. Institute on Education Statistics; Washington, DC: 2007. http://nces.ed.gov/nationsreportcard

Nagy W, Berninger V, Abbott R. Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. Journal of Educational Psychology. 2006; 98:134–147. doi:10.1037/0022-0663.98.1.134.

Nagy, WE.; Scott, J. Vocabulary processes.. In: Kamil, M.; Mosenthal, PB.; Pearson, PD.; Barr, R., editors. Handbook of Reading Research. Vol. III. Erlbaum; Mahwah, NJ: 2000. p. 269-284.

O'Connor RE, Bell K, Harty K, Larkin L, Sackor S, Zigmond N. Teaching reading to poor readers in the intermediate grades: A comparison of text difficulty. Journal of Educational Psychology. 2002; 94:474–485. doi:10.1037//0022-0663.94.3.474.

O'Connor RE, Harty KR, Fulmer D. Tiers of intervention in kindergarten through third grade. Journal of Learning Disabilities. 2005; 38(6):532–538. doi:10.1177/00222194050380060901. [PubMed: 16392695]

O'Connor R, White A, Swanson HL. Repeated reading versus continuous reading: Influences on reading fluency and comprehension. Exceptional Children. 2007; 74:31–46.

O'Sullivan, R. Protecting sea turtles. National Geographic School Publishing; Washington, DC: 2003.

Pinnell, GS.; Pikulski, JJ.; Wixson, KK.; Campbell, JR.; Gough, PB.; Beatty, AS. Listening to children read aloud. Office of Educational Research and Improvement, U.S. Department of Education; Washington, DC: 1995.

Raphael TE, Au KH. QAR: Enhancing comprehension and test taking across grades and content areas. Reading Teacher. 2005; 59:206–221. doi:10.1598/RT.59.3.1.

Rapp DN, van den Broek P, McMaster KL, Kendeou P, Espin CA. Higher-order comprehension processes in struggling readers: A perspective for research and intervention. Scientific Studies of Reading. 2007; 11:289–312.

Rosenshine B, Meister C. Reciprocal teaching: A review of the research. Review of Educational Research. 1994; 64(4):479–530. doi:10.2307/1170585.

Simmons D, Hairrell A, Edmonds M, Vaughn S, Larsen R, Willson V, Rupley W, Byrns G. A comparison of multiple-strategy methods: Effects on fourth-grade students' general and content-specific reading comprehension and vocabulary development. Journal of Research on Educational Effectiveness. 2010; 3:121–156. doi:10.1080/19345741003596890.

Singer JD. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics. 1998; 24(4):322–354. doi: 10.3102/10769986023004323.

Snow, C., editor. Reading for understanding: Toward a R & D program in reading comprehension. RAND Education; Santa Monica, CA: 2002. www.rand.org/pubs/monograph_reports/MR1465/MR1465.pdf

Speece DL, Ritchey KD, Silverman RD, Schatschneider C, Walker C, Andrusik K. Identifying children in middle childhood who are at risk for reading problems. School Psychology Review. 2010; 39:258–276. http://www.nasponline.org. [PubMed: 21472039]

Stage SA, Abbott RD, Jenkins JR, Berninger VW. Predicting response to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal IQ-word reading discrepancy: Failure to validate discrepancy method. Journal of Learning Disabilities. 2003; 36:24–33. doi:10.1177/00222194030360010401. [PubMed: 15490889]

Therrien WJ, Hughes C. Comparison of repeated reading and question generation on students' reading fluency and comprehension. Learning Disabilities: A Contemporary Journal. 2008; 6(1):1–16.

Therrien WJ, Wickstrom K, Jones K. Effect of a combined repeated reading and question generation intervention on reading achievement. Learning Disabilities Research & Practice. 2006; 21(2):89–97. doi:10.1111/j.1540-5826.2006.00209.x.

Torgesen, JK.; Wagner, RK.; Rashotte, CA. Test of Word Reading Efficiency: Examiner's manual. Pro-Ed; Austin, TX: 1999.

Vaughn S, Linan-Thompson S, Hickman P. Response to instruction as a means of identifying students with reading/learning disabilities. Exceptional Children. 2003; 69:391–409.

Wanzek J, Wexler J, Vaughn S, Ciullo S. Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. Reading and Writing. 2010; 23:889–912. doi:10.1007/s11145-009-9179-5. [PubMed: 21072128]

Woodcock, R.; McGrew, K.; Mather, N.; Schrank, F. Woodcock-Johnson Tests of Achievement. Third edition. Riverside Publishing; Itasca, IL: 2001.

Zeno, SM.; Ivens, SH.; Millard, RT.; Duvuri, R. The educator's word frequency guide. Touchstone Applied Science Associates, Inc.; Brewster, NY: 1995.

**Table 1**

Participant Demographics

| | Screening Sample (*n* = 463) | | Intervention Group (*n* = 57) | | Control Group (*n* = 66) | |
|---|---|---|---|---|---|---|
| | ***n*** | **%** | ***n*** | **%** | ***n*** | **%** |
| Gender | | | | | | |
| Male | 228 | 49.2 | 34 | 59.6 | 39 | 59.1 |
| Female | 235 | 50.8 | 23 | 38.9 | 27 | 40.9 |
| Race | | | | | | |
| African American | 107 | 23.7 | 25 | 44.6 | 28 | 43.8 |
| Caucasian | 304 | 67.3 | 26 | 46.4 | 34 | 53.1 |
| Biracial or Other | 41 | 9.1 | 5 | 8.9 | 2 | 3.1 |
| Mother's Education | | | | | | |
| < High School Grad | 6 | 1.4 | 1 | 1.8 | 2 | 3.1 |
| High School Graduate | 242 | 54.9 | 39 | 70.9 | 46 | 71.9 |
| College Degree | 97 | 22.0 | 10 | 18.2 | 10 | 15.6 |
| Graduate Degree | 95 | 21.8 | 5 | 9.1 | 6 | 9.4 |

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

**Table 2**

Descriptive Statistics

| | Intervention | | Control | |
|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* |
| Initial Predicted Probability | 0.78 | 0.19 | 0.76 | 0.19 |
| Pretest Measures | | | | |
| CBM Maze | 4.79 | 1.91 | 4.64 | 1.78 |
| GMRT Reading Comprehension | 86.72 | 6.59 | 86.39 | 7.10 |
| TOWRE Phonemic Decoding Efficiency | 94.96 | 12.12 | 93.55 | 10.16 |
| TOWRE Sight Word Efficiency | 98.78 | 11.01 | 98.63 | 8.59 |
| WJ III Word Attack | 96.76 | 9.12 | 96.40 | 8.48 |
| WJ III Word Identification | 94.29 | 8.65 | 94.80 | 7.55 |
| Progress Monitoring | | | | |
| CBM Passage Reading Fluency 1 | 92.25 | 23.69 | 93.87 | 21.89 |
| CBM Passage Reading Fluency 2 | 103.70 | 24.24 | 105.87 | 25.56 |
| CBM Passage Reading Fluency 3 | 104.63 | 24.20 | 108.13 | 25.90 |
| CBM Passage Reading Fluency 4 | 109.20 | 26.35 | 109.46 | 26.47 |
| CBM Word Identification Fluency 1 | 55.18 | 14.61 | 55.91 | 12.26 |
| CBM Word Identification Fluency 2 | 58.65 | 15.57 | 59.18 | 12.66 |
| CBM Word Identification Fluency 3 | 58.04 | 15.04 | 59.30 | 12.77 |
| CBM Word Identification Fluency 4 | 60.35 | 15.95 | 60.02 | 12.86 |
| Posttest | | | | |
| CBM Maze | 9.00 | 4.03 | 8.44 | 3.85 |
| GMRT Reading Comprehension | 86.70 | 8.44 | 87.52 | 8.82 |
| TOWRE Phonemic Decoding Efficiency | 98.09 | 13.96 | 95.83 | 11.78 |
| TOWRE Sight Word Efficiency | 97.88 | 12.40 | 93.55 | 10.16 |
| WJ III Word Attack | 97.79 | 15.61 | 98.42 | 9.51 |
| WJ III Word Identification | 97.19 | 6.99 | 96.59 | 8.44 |
| ASKIT Comprehension | 24.25 | 5.02 | 21.44 | 5.03 |
| ASKIT Reading Accuracy | 0.97 | 0.02 | 0.96 | 0.04 |
| Science Knowledge | 17.63 | 4.40 | 14.92 | 4.02 |
| Posterior Predicted Probability | 0.57 | 0.06 | 0.69 | .05 |
| Moderators | | | | |
| SSRS Academic Competence | 88.26 | 7.95 | 87.34 | 8.85 |
| TRR Overall Rating | 2.23 | 0.85 | 2.32 | 0.83 |
| TRR (Number of Reading Problems) | 2.23 | 1.91 | 2.17 | 1.96 |
| Sum of Extra Services | 0.96 | 1.04 | 1.05 | 1.22 |
| Intervention Attendance (24 sessions) | 22.16 | 2.09 | - | - |
| Tutor Rating of Attention | 4.03 | 0.63 | - | - |

*Note*. ASKIT = Assessment of Knowledge and Strategy Use for Information Text; CBM = Curriculum Based Measurement; GMRT = Gates MacGinitie Reading Test; SSRS = Social Skills Rating System; TRR = Teacher Reading Rating; TOWRE= Test of Word Reading Efficiency; TOSWRF = Test of Silent Word Reading Fluency; WJ III = Woodcock Johnson Test of Achievement, Third Edition. CBM measures are averaged across two probes and converted to items per minute. Standard scores (mean =100, *SD* = 15) are presented for norm referenced tests

**Table 3**

Random Effects ANCOVA/ANOVA Results for Pretest-Posttest Variables

| | df | F | p |
|---|---|---|---|
| CBM Maze | | | |
| Group | 1, 16 | 0.56 | .466 |
| Pretest | 1, 103 | 12.85 | <.001 |
| GMRT Reading Comprehension | | | |
| Group | 1, 16 | 0.67 | .424 |
| Pretest | 1, 103 | 32.30 | <.001 |
| TOWRE Phonemic Decoding Efficiency | | | |
| Group | 1, 15 | 0.00 | .983 |
| Pretest | 1, 101 | 400.88 | <.001 |
| TOWRE Sight Word Efficiency | | | |
| Group | 1, 15 | .64 | .437 |
| Pretest | 1, 101 | 160.78 | <.001 |
| WJ III Word Attack | | | |
| Group | 1, 15 | 0.08 | .775 |
| Pretest | 1, 101 | 175.39 | <.001 |
| WJ III Word Identification | | | |
| Group | 1, 15 | 0.29 | .596 |
| Pretest | 1, 101 | 362.81 | <.001 |
| ASKIT Comprehension | | | |
| Group | 1, 16 | 10.09 | .006 |
| Science Assessment | | | |
| Group | 1, 16 | 12.70 | .003 |
| ASKIT Comprehension: Extra Services Moderator | | | |
| Group | 1, 16 | 0.87 | 0.363 |
| Extra Services | 1, 101 | 2.66 | 0.106 |
| Group × Extra Services | 1, 101 | 5.49 | 0.021 |

*Note*. ASKIT = Assessment of Knowledge and Strategy Use for Information Text; CBM = Curriculum Based Measurement; GMRT = Gates MacGinitie Reading Test; TOWRE= Test of Word Reading Efficiency; TOSWRF = Test of Silent Word Reading Fluency; WJ III = Woodcock Johnson Test of Achievement, Third Edition.

## Table 4

Conditional Growth Models for Passage Reading Fluency and Word Identification Fluency

| | df | F | p |
|---|---|---|---|
| Passage Reading Fluency | | | |
| Pretest | 1, 586 | 1786.21 | <0.001 |
| Group | 1, 16 | 0.07 | 0.795 |
| Time | 1, 586 | 180.60 | <0.001 |
| Group × Time | 1, 586 | 1.37 | 0.242 |
| Word Identification Fluency | | | |
| Pretest | 1, 584 | 2514.54 | <0.001 |
| Time | 1, 584 | 0.91 | 0.340 |
| Time × Time | 1, 584 | 42.61 | <0.001 |
| Group | 1, 16 | 0.78 | 0.390 |
| Time × Group | 1, 584 | 2.87 | 0.091 |
| Time × Time × Group | 1, 584 | 3.97 | 0.468 |
| Passage Reading Fluency: Interaction with Initial Predicted Probability | | | |
| Pretest | 1, 582 | 1130.24 | <0.001 |
| Time | 1, 582 | 207.67 | <0.001 |
| Group | 1, 16 | 0.06 | 0.805 |
| Group × Time | 1, 582 | 1.34 | 0.247 |
| Initial Predicted Probability | 1, 582 | 5.78 | 0.016 |
| Initial Predicted Probability × Group | 1, 582 | 4.02 | 0.045 |
| Initial Predicted Probability × Group × Time | 2, 582 | 3.37 | 0.035 |