Published in final edited form as:
J Learn Disabil. 2012 ; 45(6): 515–525. doi:10.1177/0022219411402692.

Effects of Intensive Reading Intervention for Eighth-Grade Students With Persistently Inadequate Response to Intervention

Sharon Vaughn, PhD[1], **Jade Wexler, PhD**[1], **Audrey Leroux**[1], **Greg Roberts, PhD**[1], **Carolyn Denton, PhD**[2], **Amy Barth, PhD**[3], and **Jack Fletcher, PhD**[3]

[1]University of Texas at Austin, Austin, TX, USA

[2]University of Texas Health Science Center, Houston, TX, USA

[3]University of Houston, Houston, TX, USA

Abstract

The authors report the effects of a yearlong, very small-group, intensive reading intervention for eighth-grade students with serious reading difficulties who had demonstrated low response to intervention (RTI) in both Grades 6 and 7. At the beginning of Grade 6, a cohort of students identified as having reading difficulties were randomized to treatment or comparison conditions. Treatment group students received researcher-provided reading intervention in Grade 6, which continued in Grade 7 for those with low response to intervention; comparison students received no researcher-provided intervention. Participants in the Grade 8 study were members of the original treatment ($N = 28$) and comparison ($N = 13$) conditions who had failed to pass a state-mandated reading comprehension test in both Grades 6 and 7. In Grade 8, treatment group students received a 50-minute, daily, individualized, intensive reading intervention in groups of two to four students per teacher. The results showed that students in the treatment condition demonstrated significantly higher scores than comparison students on standardized measures of comprehension (effect size = 1.20) and word identification (effect size = 0.49), although most continued to lack grade-level proficiency in reading despite 3 years of intervention. Findings from this study provide a rationale for intensive intervention for middle school students with severe reading difficulties.

Keywords

middle school; response to intervention; intensive reading; small group

While considerable research in the past two decades has addressed many important questions about reading instruction, including the efficacy of various intervention practices for primary-grade students with reading difficulties (for reviews, see Foorman, 2003; Wanzek & Vaughn, 2009), there has been inadequate research addressing two issues of high importance. The first concerns effective practices for intervening with students who are inadequate responders to typically effective interventions (D. Fuchs, Fuchs, & Stecker, 2010; Wanzek & Vaughn, 2007), and the second is the extent to which interventions can effectively improve reading comprehension for older students (after Grade 3) with significant and persistent reading difficulties (Deshler & Hock, 2007; Vaughn & Fletcher,

Corresponding Author: Sharon Vaughn, The Meadows Center for Preventing Educational Risk, The University of Texas at Austin, 1 University Station, D4900, Austin, Texas 78712-0365 srvaughn@austin.utexas.edu.

Declaration of Conflicting Interests The authors declared no potential conflict of interest with respect to the authorship and/or publication of this article.

2010; Wanzek, Vaughn, Roberts, & Fletcher, in press). We briefly summarize the research on these two important issues and then describe the background for a study that examined the effects of intensive intervention provided to students in Grade 8 who had demonstrated inadequate response to intervention (RTI) for the previous two years.

Low Response to Intervention

One of the primary mechanisms for determining low RTI is to provide students with reading difficulties an intervention that is associated with improved outcomes and then identify as inadequate responders those students who do not meet preidentified criteria. Our understanding of this RTI approach was influenced by Vellutino and colleagues (1996), who identified participants as making low growth or very low growth after they were in a generally effective intervention. Vellutino provided additional reading intervention to students who demonstrated low growth (Vellutino, Scanlon, & Lyon, 2000), describing how we might think about RTI in reading. Though largely focused on elementary students, RTI has increasingly been implemented in middle and high schools as a means of providing increasingly intensive interventions for students with reading difficulties (Fletcher & Vaughn, 2009). There have been a few studies that have examined the effects of multitiered interventions by identifying participants as inadequate responders to an initial intervention and then providing the inadequate responders with either additional intervention or a different intervention. Most of these multitier intervention studies were conducted with students at risk for reading difficulties in the primary grades (Al Otaiba, 2000; Berninger et al., 2002; Vadasy, Sanders, Peyton, & Jenkins, 2002; Vaughn, Linan-Thompson, & Hickman, 2003; Vellutino et al., 1996).

Wanzek (2005) extended a synthesis on student characteristics associated with low RTI conducted by Al Otaiba and Fuchs (2002) identifying 42 studies related to response to intervention for elementary students. Overall findings revealed that rapid naming ability, phonological awareness, and verbal ability differentiated students with adequate and inadequate RTI. Wanzek's synthesis also indicated that RTI could be enhanced by improving classroom instruction—typically described as Tier 1 instruction (Center, Freeman, & Robertson, 2001; Juel & Minden-Cupp, 2000). As expected, the synthesis (Wanzek, 2005) also indicated that students with disabilities (e.g., language impairments and reading disabilities) were more likely to demonstrate insufficient response to interventions than students without disabilities (Fazio, 1997; Hurford, 1990).

Secondary Students and Response to Intervention

In 2007, the National Assessment of Educational Progress reported that 69% of eighth-grade students were unable to successfully derive meaning from grade-level text, which is perhaps the most critical skill for students whose primary responsibility is to learn from text in content-area classes (Biancarosa & Snow, 2004; Perfetti, Landi, & Oakhill, 2005). With such a high prevalence of reading problems continuing into the middle grades, and an increasing focus on improving high school retention and preparing students for postsecondary learning, adolescent reading instruction has become increasingly important (Kamil et al., 2008). However, there are few studies evaluating the effects of secondary (Tier 2) interventions and even fewer evaluating the effects of tertiary (Tier 3) interventions for students in secondary grades.

Two syntheses have addressed the impact of supplemental reading interventions for students in Grades 4 through 12. Edmonds et al. (2009) conducted a meta-analysis examining the effects on reading comprehension outcomes of interventions that provided instruction in decoding, fluency, vocabulary, or comprehension for adolescents with reading difficulties. A total of 13 studies met the criteria for a meta-analysis (e.g., comparison group), yielding a

mean weighted effect size in the moderate range in favor of treatment students over comparison students. Promising interventions provided targeted reading intervention in comprehension, multiple reading components, or word recognition strategies.

Scammacca et al. (2007) provided further evidence for positive outcomes related to adolescent reading interventions providing separate findings for students with learning disabilities. In a meta-analytic synthesis, single- and multi-component interventions were examined to determine the effects of interventions that targeted word study, fluency, vocabulary, and comprehension on various reading-related outcomes. Scammacca et al. located 31 studies that met criteria for review and coded them on the following variables: (a) intervention characteristic (e.g., word study, vocabulary, fluency, etc.), (b) participant age/grade, and (c) intervention instructor (e.g., teacher, researcher, paraprofessional, etc.). The majority of studies included in the synthesis reported outcomes using researcher-developed measures, which have been associated with higher effect sizes in numerous syntheses (Elleman, Lindo, Morphy, & Compton, 2009). For example, in the Elleman et al. (2009) synthesis of vocabulary interventions, effect sizes decreased on average from .50 for researcher-developed measures to .10 for normative measures. When the 11 studies using only standardized measures were examined, the average effect size was .42. Comprehension and vocabulary practices were associated with the highest outcomes, and word study was associated with moderate effects. It is important to interpret these effects with caution as they primarily reflected researcher-developed measures. Using only normative measures (not available for vocabulary), the impact of moderator variables shifted, with word study and comprehension strategies demonstrating the highest effect sizes. Higher effect sizes were also associated with researcher-implemented interventions and middle school as opposed to high school participants. None of these studies focused on students with demonstrable inadequate response to previous interventions. Furthermore, none of these studies provided a multi-component, comprehensive approach to remediating reading difficulties for secondary students with significant and persistent reading difficulties.

Study Background

To address the need to better understand (a) the effects of intensive interventions on outcomes for secondary students with reading difficulties and disabilities and (b) the effects of reading instruction provided to secondary students within an RTI framework, we designed a series of studies with middle school students who were at risk for reading difficulties. This article reports findings from the 3rd year of a 3-year longitudinal study conducted with a single cohort of students in Grades 6, 7, and 8. In Grade 6, middle school students with reading difficulties were identified based on low performance on a state-mandated reading comprehension test and randomized to treatment or business-as-usual comparison conditions. During this year, students in both groups received enhanced Tier 1 instruction through the application of explicit instructional routines designed to support vocabulary and comprehension in content-area classes. Treatment group students received intervention in 50-minute daily sessions provided by the research team for each year, while comparison group students received no researcher-provided intervention. Findings from Year 1 showed treated students with small gains on measures of decoding, fluency, and comprehension over the course of the year (median $d = +0.16$; Vaughn et al., 2010). When interpreting effects, it is important to note that all students participated in an enhanced Tier 1 instruction and many students in the comparison condition were provided school-level interventions.

After one full year of treatment, inadequate responders were identified, again using a benchmark on the state reading accountability test. Treatment group students who did not meet the benchmark were randomized to receive either a standardized or individualized

treatment and provided a second year of intervention in Grade 7. Findings from the Year 2 study did not demonstrate significant differences between the standardized or individualized treatment, although treatment students performed significantly better than comparison on reading comprehension and outperformed comparison students on reading outcomes (median $d = 0.23$; Vaughn et al., in press).

Following the second year of intervention, inadequate responders were again identified; those in the treatment group received a third year of increasingly intensive intervention in Grade 8, while comparison group students again received whatever instruction was typically provided in their schools. This article reports on this third year of the study, extending previously reported research by addressing the following question: What are the effects of an intensive instructional intervention on the reading outcomes of students with significant reading difficulties who had demonstrated very low growth in 2 years of previous small-group reading intervention, relative to low growth students in a comparison condition, who received whatever instruction was typically provided to low-performing readers in their schools?

Method

Design

The general design is a multigated, longitudinal, randomized trial with increasingly intense tiers of intervention. At-risk sixth-graders in Year 1 of the study who remained at risk until eighth grade were included in the sample for this study. Findings are reported from Year 3, when the targeted cohort was in eighth grade. While the comparison group was determined using the same end-of-year inadequate response criteria used in identifying treatment participants, and as such represent the best of available counterfactuals, treatment condition students in Year 3 of the study had the benefit of two prior years of intervention. All students in special education in both treatment and comparison conditions continued to receive all special education services as reflected in their Individualized Education Programs (IEPs).

Participants

The research was conducted with institutional review board approval in highly diverse middle schools in two large urban cities (one large district, one medium district) in the southwestern United States, with approximately half the sample from each site. School populations ranged in size from 498 to 1,145. In 2006–2007 (Year 1) we selected all struggling readers using the state accountability test results (*Texas Assessment of Knowledge and Skills* [TAKS]) to identify struggling readers with a scale score that approximates the 30th percentile on other norm-referenced reading comprehension assessments (TAKS below 2,150). Students exempted from the TAKS due to special education status because of very low reading achievement were also included. Students were randomized with two treatment students for every one comparison student because (a) in a response to intervention model we expected considerable attrition and (b) for Year 2 of the study we examined effects of two treatments (standardized and individualized), and the design required a larger sample of treatment students.

An intent-to-treat model informed data management and analysis. For the Year 2 study, students in the treatment condition who had demonstrated inadequate instructional response in Year 1 were randomized after posttesting in the spring of Year 1 (2007) to individualized or standardized treatment conditions. All eligible cases ($n = 182$) were included in the analysis sample. Of the 182 sampled cases, 150 returned to participating schools in the fall of 2007, 38 comparison students, 55 in the individualized condition, and 57 in the standard

protocol condition. In the spring of 2008, there were 36 comparison students, 51 students in the individualized condition, and 46 in the standard protocol condition. Pretest scores for students not continuing through spring of 2008 were compared across the three groups. There were no statistically significant differences. Over 90% of the coverage estimates (amount of data present in each cell of the measure by occasion matrix) were at or above .75 across all outcomes and measurement occasions.

Students who received treatment in Year 3 (2008–2009) were those who (a) participated in intervention in Year 1 of the study (Vaughn et al., 2010), (b) participated in an additional intervention in Year 2 (either standardized or individualized protocol; Vaughn et al., in press), *and* (c) failed to meet the criteria for adequate RTI in both Years 1 and 2. In Year 1, students were considered low responders to intervention if they scored less than 2,100 scaled scored points on the *Texas Assessment of Knowledge and Skills*, *or* scored less than 90 standard scored points on the *Woodcock-Johnson III* (WJ-III) *Letter-Word Identification* assessment at posttest, *or* scored less than 90 standard scored points on the *Group Reading Assessment and Diagnostic Evaluation* (GRADE) *Comprehension Composite* assessment at posttest. The same criteria were used to determine nonresponse at the end of Year 2, applying a cut score of less than 2,150 scaled scored points on the TAKS (one half of one standard error of measurement above the state-established passing score, in order to identify all students who may be at risk, reducing the number of false negatives), instead of 2,100. In all, 28 students participated in the Year 3 treatment. Students in the comparison group were those who had been randomly assigned to the comparison condition in Year 1 of the study *and* met the same end-of-year criteria for inadequate response each year as treated students. Of the 28 treatment students and 13 comparisons followed for all 3 years of the intervention, 92% received free or reduced lunch. Of the treatment participants, 69% were male and 39% were identified as having limited English proficiency, with 61% Hispanic, 31% African American, and 8% Caucasian. Of the treatment students, 35% were identified as receiving special education services (1 identified as emotionally disturbed, 7 as having learning disabilities, and 1 speech and language). For the comparison students, 43% were male, 15% were identified as having limited English proficiency, with 61% African American, 36% Hispanic, and 3% Caucasian. In addition, 15% ($n = 2$) of the comparison students were identified as receiving special education, both with learning disabilities. The age range of students was 13.09 to 15.87 for treatment students with a mean of 14.18, and for comparisons, the range was 13.82 to 14.98 for a mean of 14.40.

Teacher participants—The researchers hired two female intervention teachers, both with master's degrees in education with high levels of knowledge and experience (i.e., 6 and 15 years) in teaching students with significant reading disabilities. The research team provided the intervention teachers with approximately 60 hours of professional development prior to teaching. They also participated in biweekly staff development meetings with ongoing on-site feedback and coaching (once every 1 to 2 weeks) throughout the year.

Description of Intervention

Students in the treatment condition were provided a 50-minute reading intervention class during an elective period. Students were taught in groups of two to four. Comparison students participated in an elective. Electives ranged from music, art, a foreign language, and physical education (i.e., football). The focus of the treatment group instruction was a clinical teaching model that used data to inform decision making (see Vaughn et al., in press, for a further description). Teachers tailored instruction to meet students' individual needs, initially using students' Year 2 posttest scores and then teacher-created curriculum based measures (CBM) progress monitoring data throughout the year. Based on students' diagnostic assessments and progress-monitoring data, teachers organized lessons based on

individual student performance. Lessons were designed and implemented to address students' needs in phonics, word reading, fluency, vocabulary, and comprehension. For example, using diagnostic data from the *Woodcock-Johnson III* administered at Year 2 posttest, students with a standard score of 95 or above on the *Word Attack* subtest were provided minimal instruction on word study and maximum instruction on text reading with emphasis on comprehension and vocabulary development. Students who scored below a 95 standard score on the *Word Attack* subtest were provided a more intensive focus on word study instruction as well as vocabulary and comprehension strategy instruction.

Teachers were provided with requirements for use of instructional time. Students with adequate decoding skills received 35 to 45 minutes of instruction in vocabulary and morphology, 170 to 180 minutes of comprehension and text reading, and 15 to 25 minutes of a motivational component during each 5-day week. Students with below average word reading scores were provided 100 to 110 minutes of intensive word study and text reading instruction, 35 to 45 minutes of vocabulary and morphology instruction, 70 to 80 minutes of comprehension and text reading instruction, and 15 to 25 minutes of a motivational component each week.

Teachers began with these relative emphases for each student and then adjusted them based on biweekly progress monitoring. Specifically, teachers developed CBMs to reflect skills within each reading component they had addressed during the previous 2-week period. For example, if the target of word study for the week was breaking multisyllable words into parts and then reading the parts and blending them into a word, the CBM for the week would provide a list of practiced words to determine how well each student could complete the task. Students were also asked to read connected text at increasingly difficult levels and to answer questions and provide summaries of what they read. Teachers used the results to determine if a student had mastered, was still emerging, or was deficient on each skill and adjusted instruction accordingly with guidance from the research team. These CBMs were designed and used for diagnostic purposes only. More formal decisions regarding student progress were based on progress monitoring checks.

Overall, teachers followed a similar scope and sequence of research-based comprehension strategy instruction (e.g., strategies for finding main ideas and summarizing text) for all students, but had access to a variety of instructional materials and could modify pacing and materials in response to students' needs. Using a variety of narrative and expository text, teachers taught and scaffolded use of the strategies before, during, and after reading. Before reading, students were taught to preview passages and make predictions, and teachers focused on teaching essential vocabulary related to passage comprehension using approaches such as semantic feature analysis for developing understanding. Teachers and students participated in discussions about definitions, characteristics, and relationships of words. Students were also taught strategies for self-monitoring reading comprehension and strategies to repair their misunderstandings. An emphasis was placed on self-monitoring and “repairing” comprehension at the sentence level. Students learned to generate main ideas for paragraphs and eventually multiparagraph text, as well as generating summaries and wrap-up statements. Teachers used information text on science- and social studies–related topics at varying levels, a variety of novels, as well as expository text that students would encounter on a regular basis, such as text taken from their own textbooks. Word study instruction was also flexible, but teachers primarily followed a sequential scope and sequence and used lessons from an explicit, intensive multisensory word study program (Wilson, 1996). Teachers progressed through instruction in the Wilson (1996) program in a sequential manner, but had the flexibility of varying pacing and lesson implementation according to student need. Finally, a motivation component was built into the daily individualized lessons, with weekly expectations for purposeful and motivational text

selection, student and teacher goal setting, evaluation conferences, and positive phone calls home.

Fidelity Data

Two project directors, one from each site, collected fidelity data four times each year for each teacher providing intervention. Interrater reliability between the two observers was calculated as the number of agreements divided by the sum of the number of agreements and number of disagreements. Interrater reliability was 94%. These same raters were reliable during the previous year on a study using a similar treatment (Vaughn, et al., in press).

The goal was to establish adherence to the treatment model and the extent to which the intervention was provided as intended. Fidelity was conceptualized as the difference between the intended (or normative) program model and the implemented model (Chen, 1990). The normative model for the intervention was represented by teachers following specified program elements/activities (i.e., word study, comprehension, vocabulary instruction) within specified time amounts within a week. Fidelity was collected using a 3-point Likert-type rating scale ranging from 1 (*low*) to 3 (*high*) to assess the extent to which the teacher completed each required instructional element or activity and the overall quality of implementation, which included active student engagement. The individualized program model was conceptualized to respond to the "individualized" intent of the treatment, asking teachers to teach particular reading components (i.e., word study/text reading, vocabulary/ morphology) for a set time on a weekly basis according to student needs. Ratings of "overall quality" of implementation for each specified reading component were also collected.

Individual teachers' mean implementation scores for the intervention ranged from 2.03 to 2.63, with a group average of 2.33. Mean quality scores for the intervention ranged from 2.50 to 2.56, with a group average of 2.53. The mean total fidelity ranking, including implementation and quality ratings, ranged from 2.26 to 2.56 for the intervention. Copies of the fidelity protocols can be obtained at www.meadowscenter.org.

Measures

All measures, except the *Texas Assessment of Knowledge and Skills* (Texas Education Agency, 2004), were administered by members of the research team who were hired and trained by an assessment coordinator and were blind to whether students were in treatment or comparison conditions. The *Texas Assessment of Knowledge and Skills* (Texas Education Agency, 2004) is a criterion-referenced reading comprehension test that is the Texas accountability test. The TAKS is not timed and includes different assessments used for each grade that are criterion-referenced and aligned with grade-based standards from the Texas Essential Knowledge and Skills (TEKS). The internal consistency (coefficient alpha) of the Grade 7 test is .89 (Texas Education Agency, 2004). We used it as an initial screening assessment and then a benchmark assessment because it is reliable, represents an accountability "high-stakes" assessment implemented in all states, and has good construct validity as a measure of reading comprehension. In preliminary latent-variable analyses of the students in Grades 6 through 8, the TAKS measure loaded strongly on the WJ-III *Passage Comprehension* subtest and the *Group Reading Assessment and Diagnostic Evaluation* (Williams, 2001).

All of the following measures were obtained at pre- and posttest.

Decoding and spelling—We assessed word reading accuracy for real words and pseudowords with the *Letter-Word Identification* and *Word Attack* subtests of the *Woodcock-Johnson III Tests of Achievement* (Woodcock, McGrew, & Mather, 2001).

Coefficient alphas based on a sample from the previous year of 327 struggling readers and 249 typical readers who contributed data throughout the year for *Letter-Word Identification* and *Word Attack* ranged from .93 to .97.

Fluency—The *Sight Word Efficiency* and *Phonemic Decoding Efficiency* subtests from the *Test of Word Reading Efficiency* (TOWRE; Torgesen, Wagner, & Rashotte, 1999) assessed word list fluency for real words and pseudowords. Internal consistency for different forms of this well-standardized test exceeds .90.

The Test of Silent Reading Efficiency and Comprehension (TOSREC; Wagner, Torgesen, Rashotte, & Pearson, 2010) is a 3-minute, group-based assessment that assesses silent reading of connected text for comprehension. Students are presented with a series of short sentences and are required to read and verify the truthfulness of as many sentences as possible. The raw score is the number of sentences correctly identified as true or false, minus the number of incorrect responses, within the time limit; skipped sentences are ignored, and if the number of incorrect responses exceeds the number correct, a raw score of 0 is recorded. The TOSREC was standardized on 2,000 students from Grades 4 through 9. The standard score was the dependent measure utilized. The mean intercorrelation of performances across the five time points in the Grade 6 sample of 327 struggling readers and 249 typical readers was .79 for standard scores and .80 for raw scores. The mean intercorrelation of performances across the five time points in the Grade 7–8 sample of 436 struggling readers and 440 typical readers was .96 for standard scores and .96 for raw scores. These correlations likely underestimate reliability since some students received intervention and may have changed their rank order over time. The criterion related validity of TOSREC with TAKS *Reading* was 0.56 in a sample of 1,421 middle school students in Grades 6 through 8.

The *AIMSweb Reading Maze* (M. R. Shinn & Shinn, 2002), a 3-minute, group-administered curriculum-based assessment of fluency and comprehension, was administered at all five time points. Students are presented with a 150- to 400-word passage, and for every seventh word after the first sentence students are asked to identify a correct target word from among three choices. The raw score is the number of targets correctly identified within the time limit, and was the dependent measure utilized. *AIMSweb* provides 15 different stories for sixth grade, and the particular story any individual student received was randomly determined within school and treatment group. These measures were not equated, although stories were chosen based on reading level. Reliability and validity information is reported (L. S. Fuches & Fuchs, 1992 Jenkins & Jewell, 1993; J. Shinn, Deno, & Espin, 2000; M. R. Shinn & Shinn, 2002). The mean intercorrelation of performances across the five time points in the Grade 6 sample of 327 struggling readers and 249 typical readers was .79. The mean intercorrelation of performances across the five time points in the Grades 7–8 sample of 436 struggling readers and 440 typical readers was .95. Mean intercorrelations of performance are likely to be an underestimate of reliability since some students received intervention and may have changed their rank order over time. The criterion-related validity of *AIMSweb Maze Reading* with TAKs *Reading* was 0.39 in a sample of 1,421 middle school students in Grades 6 through 8.

Comprehension—The WJ-III *Passage Comprehension* subtest is an individually administered cloze-based assessment in which students read a passage and fill in a missing word. Coefficient alphas in the entire sample of 327 struggling readers and 249 typical readers were .94 at pretest and .85 at posttest.

The *Gates-MacGinitie Reading Test* was administered as a timed, group-administered assessment of reading comprehension consisting of expository and narrative passages

ranging in length from 3 to 15 sentences. After reading each passage silently, students then answer three to six multiple-choice questions related to the most recently read passage. Items increase in difficulty as the student progresses through the test during the 35-minute time limit. Internal consistency reliability ranges from .91 to .93 and alternate form reliability is reported as .80 to .87.

Results

Descriptive statistics are presented for the treatment group in Table 1 and for the comparison group in Table 2 for pretest and posttest. Absolute differences in pretest and posttest means for each measure and for each condition are provided; differences in pretest and posttest performance are also expressed in standardized units (Hedge's *g* for repeated measures using the small sample correction) for each group. Measures are organized according to domain. Domains include word reading accuracy, fluency, and comprehension. For the treatment group, there is a general pattern of improvement on all measures from pretest for posttest, with the exception of the WJ-III *Letter-Word* subtest, *AIMSweb Mazes* measure, and the *Gates-MacGinitie Passage Comprehension* subtest. In the comparison group, means at posttest are lower than pretest means. Year 1 pretest scores from the treatment and comparison groups were tested for differences, and significant differences between the two groups were not found after the Benjamini-Hochberg Correction was used, indicating that the two groups were similar on all measures at the beginning of the 3-year study and at the beginning of this study.

To assist readers in examining findings for treatment and comparison students with disabilities, primarily learning disabilities, we provide the range of findings for measures separately for these students in Table 3. However, it is important to remember that the sample sizes are small with nine in treatment condition and only two in comparison condition. Results indicate that students in the sample identified with disabilities (predominantly learning disabilities) scored lower than students without disabilities on all assessments; however, students identified with disabilities in the treatment group consistently outscored those identified with learning disabilities in the comparison group.

Analysis of covariance (ANCOVA) was used to test for group differences. In randomized designs, ANCOVA is used to minimize error variance in the dependent variable, increasing the precision of the estimate and the power of the analysis. The pretest score was used as the covariate. Assumptions regarding homogeneity of regression were met. Table 4 presents the ANCOVA results. There were statistically significant differences between treatment and comparison groups on the *Gates-MacGinitie* assessment, $F(1, 37) = 12.475$, $p = .001$, with adjusted posttest means of 82.78 and 74.18, respectively. Differences on the WJ-III *Letter-Word Identification* subtest, $F(1, 36) = 6.564$, $p = .015$, and on the TOWRE *Phonemic Decoding* subtest, $F(1, 36) = 5.114$, $p = .030$, were also statistically significant. However, differences on TOWRE *Phonemic Decoding* did not differ significantly when Type I error rate was controlled using the Benjamini-Hochberg correction. There were no statistically significant differences on WJ-III *Word Attack*, TOSREC, TOWRE sight words, WJ-III passage comprehension, and *AIMSweb Mazes*, although in all cases except for *AIMSweb Mazes* the effect sizes were in favor of the treatment condition and were practically significant.

Discussion

This study reports findings from the final year of a 3-year intervention for students in eighth grade who were participants in a 3-year RTI treatment in which they were identified as demonstrating reading difficulties in Grade 6 and provided increasingly intensive

interventions. This article reports findings from an intensive, individualized reading intervention on the reading outcomes of students with significant reading difficulties and disabilities who had responded with very low growth to 2 years of previous intervention and compares their results with low growth students in a comparison condition. Interpretation of these findings should be considered in light of the intractable reading difficulties exhibited by the target students.

Eighth-grade students who had been consistently inadequate responders to researcher-provided treatment over 2 previous years (Vaughn et al., 2010, in press) demonstrated significantly higher scores at posttest than comparison students on the *Gates-MacGinitie Test of Reading Comprehension* and the *Woodcock-Johnson Word Identification Test*. While students in the treatment condition outperformed comparison students, posttest standard scores indicate that they did not close the gap with typically performing peers. Instead, the benefit of continued remedial instruction largely reflected the continuing downward trend of students in the comparison condition compared to the relatively steady performance of the students in the treatment group. The findings suggest that eighth-grade students with intractable reading disabilities may be able to maintain their status relative to their average-age classmates when provided with intensive, daily intervention across the school year; however, treated students remained poor readers. The finding that without continued remediation comparison students' reading performance deteriorates provides a strong rationale for continued remediation in reading even for adolescents with significant and persistent reading difficulties and disabilities. Such findings are reminiscent of other studies of reading outcome for students in special education, where they do not fall further behind, but do not exhibit accelerated progress (Hanushek, Kain, & Rivkin, 1998).

A persistent question for secondary students with reading difficulties is the extent to which reading interventions for older students should be terminated and replaced with compensatory or other instruction. This question has plagued secondary teachers of students with learning disabilities for decades (Mastropieri, Scruggs, & Graetz, 2003; Vaughn et al., in press). One of the primary arguments against continuing reading instruction is that time spent on reading might not be beneficial and may even detract from content learning. This study indicates that eighth-grade students with significant reading difficulties who are not provided intervention have significant declines in their reading performance and students provided reading interventions maintain reading skills and do not experience this same decline. We interpret the findings from this study as providing initial support for continuing reading interventions using texts that build background knowledge and understanding for content learning (e.g., science, social studies).

Because the students in this study were in the eighth grade and had significant reading difficulties, we wondered whether they would continue to be interested in learning to read and be motivated to participate in the intervention. At the end of treatment, a member of the research team who was not one of their teachers interviewed students individually to determine their views on the reading intervention and their own interest in being better readers. Confirming what we observed and were told by their reading intervention teachers, all students reported that they were interested in learning to understand what they read and enjoyed participating in the reading intervention class. One student summarized the views of many of the students by citing one of the perceived benefits of the reading class: “In other classes the teacher doesn't explain a lot and this teacher goes back when we need help.” Another student indicated his desire to continue working on becoming a better reader because of the potential positive influence being a good reader can have for one's future: “Reading will help me get my job in the future and it helps you through life, gets you where you need to go.”

This study supports the practice of providing intensive remediation for adolescents with intractable reading difficulties and disabilities. Understanding how to meet the instructional needs of such students is a serious problem challenging the design of effective interventions (Torgesen, Wagner, & Rashotte, 1997). Particularly challenging is determining the effectiveness of interventions with students identified as having the most intractable of reading problems—older students demonstrating inadequate response to intervention. This study specifically addressed these students and determined that the impact of the intervention was moderate for word reading and high for comprehension.

There is accumulating evidence that remediating reading problems in students after fourth grade will require a long-term commitment; it may be necessary to provide reading interventions throughout secondary school while also increasing instructional practices such as vocabulary and comprehension enhancements within content-area instruction (e.g., social studies, science, and math). For example, Kim, Samson, Fitzgerald, and Hartry (2010) reported findings from a randomized experiment in which students in Grades 4 through 6 participated in an afterschool intervention (modified READ 180) or a district afterschool program. Students in the treatment condition did not score significantly higher than students in the district afterschool program on measures of word reading or comprehension, though students in both groups made gains. Torgesen and colleagues (2006) reported that students in Grade 5 did not demonstrate significant gains when provided reading interventions whereas treated students in Grade 3 improved on reading accuracy, fluency, and comprehension. The authors indicated that remediating reading difficulties in students is increasingly difficult as they move up the grades. It is exceedingly challenging for middle schools since many have not hired and prepared teachers to provide intensive reading interventions to students as the teachers in this study were hired and paid by the research project.

We have been asked on numerous occasions when presenting these findings and again by the reviewers of this article to provide speculation or conjecture about what it would take to make a significant improvement in the overall reading success of students with significant reading difficulties. Perhaps if we adjusted the quality of the treatment we provided, for example, students worked more on computers or with digital text, or perhaps adjusted the breadth of reading comprehension, for example, provided students with extensive reading comprehension instruction embedded in their content-area instruction, greater improvement might be realized. It may also be a matter of when intensive interventions were provided, for example, starting in first grade rather than in sixth grade. This study does not provide answers to these questions, but they are important ones to consider for future research.

Limitations and Implications

The sample in this study is small, though not unusually so for studies of this type, particularly when considering that the sample is longitudinal and represents a cohort that participated in research for 3 years. The comparison group represents a reasonable counterfactual. However, given the longitudinal nature of the design and the fact that students were initially randomized in Grade 6, there may be some questions concerning internal validity that can be addressed by replicating this multiyear response to intervention model. The most significant implication from this study is that students with intractable reading difficulties benefit from treatments largely by maintaining reading achievement relative to their peers (i.e., based on standard scores), whereas students in the comparison group decline in reading comprehension from the beginning of the year to the end of the year, yielding an overall large treatment effect for reading comprehension. However, we think it is difficult to overstate the extensive treatment received by the target students (50 minutes a day for 3 years) and that this extensive treatment was associated with very high reading comprehension impact when contrasted with comparison students but relatively

little overall growth. Furthermore, this study does not determine whether additional time in intervention is associated with gains regardless of the intervention provided. Since reading for understanding is the hallmark of success—particularly with the increasing text demands in high school and post-secondary settings—better understanding the effects of treatment designed to enhance reading comprehension are needed.

Acknowledgments

Funding The authors disclosed receipt of the following financial support for the research and/or authorship of this article:

This research was supported by Grant P50 HD052117 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development or the National Institutes of Health.

References

Al Otaiba SD. Children who do not respond to early literacy intervention: A longitudinal study across kindergarten and first grade. Dissertation Abstracts International. 2000; 61(04):1354A. (UMI No. 9970028).

Al Otaiba S, Fuchs D. Characteristics of children who are unresponsive to early literacy intervention: A review of the literature. Remedial and Special Education. 2002; 23:300–316.

Berninger VW, Abbott RD, Vermeulen K, Ogier S, Brooksher R, Zook D, Lemos Z. Comparison of faster and slower responders to early intervention in reading: Differentiating features of their language profiles. Learning Disability Quarterly. 2002; 25:59–76.

Biancarosa, C.; Snow, CE. Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York. 2nd ed.. Alliance for Excellent Education; Washington, DC: 2004.

Center Y, Freeman L, Robertson G. The relative effect of a code-oriented and a meaning-oriented early literacy program on regular and low progress Australian students in year 1 classrooms which implement reading recovery. International Journal of Disability, Development and Education. 2001; 48:207–232.

Chen, H. Theory-driven evaluations. Sage; Newbury Park, CA: 1990.

Deshler, DD.; Hock, MF. Adolescent literacy: Where we are, where we need to go. In: Pressley, M.; Billman, AK.; Perry, KH.; Reffitt, KE.; Reynolds, JM., editors. Shaping literacy achievement: Research we have, research we need. Guilford; New York, NY: 2007. p. 98-128.

Edmonds MS, Vaughn S, Wexler J, Reutebuch CK, 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. [PubMed: 20072704]

Elleman AM, Lindo EJ, Morphy P, Compton DL. The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. Journal of Research on Educational Effectiveness. 2009; 2:1–44.

Fazio BB. Learning a new poem: Memory for connected speech and phonological awareness in low-income children with and without specific language impairment. Journal of Speech, Language, & Hearing Research. 1997; 40:1285–1297.

Fletcher JM, Vaughn S. Response to intervention: Preventing and remediating academic deficulties. Child Development Perspectives. 2009; 3:30–37. [PubMed: 21765862]

Foorman, BR. Preventing and remediating reading difficulties: Bringing science to scale. York Press; Baltimore, MD: 2003.

Fuchs D, Fuchs LS, Stecker PM. The “blurring” of special education in a new continuum of general education placements and services. Exceptional Children. 2010; 76:301–323.

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

Hanushek, EA.; Kain, JF.; Rivkin, SG. Does special education raise academic achievement for students with disabilities?. National Bureau of Economic Research; Cambridge, MA: 1998. Working Paper No. 6690

Hurford DP. Training phonemic segmentation ability with a phonemic discrimination intervention in second- and third-grade children with reading disabilities. Journal of Learning Disabilities. 1990; 23:564–569. [PubMed: 2250137]

Jenkins JR, Jewell M. Examining the validity two measures for formative teaching: Reading aloud and maze. Exceptional Children. 1993; 59:421–432.

Juel C, Minden-Cupp C. Learning to read words: Linguistic units and instructional strategies. Reading Research Quarterly. 2000; 35:458–492.

Kamil, M.; Borman, G.; Dole, J.; Kral, C.; Salinger, T.; Torgesen, J. Improving adolescent literacy: Effective classroom and intervention practices: A practice guide (NCEE 2008–4027). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education; Washington, DC: 2008.

Kim JS, Samson JF, Fitzgerald R, Hartry A. A randomized experiment of a mixed-methods literacy intervention for struggling readers in grades 4–6: Effects on word reading efficiency, reading comprehension and vocabulary, and oral reading fluency. Reading and Writing: An Interdisciplinary Journal. 2010; 23:1109–1129.

Mastropieri M, Scruggs TE, Graetz JE. Reading comprehension instruction for secondary students: Challenges for struggling students and teachers. Learning Disabilities Quarterly. 2003; 26:103–116.

Perfetti, CA.; Landi, N.; Oakhill, J. The acquisition of reading comprehension skills. In: Snowling, MJ.; Hulme, C., editors. The science of reading: A handbook. Blackwell; Oxford, England: 2005. p. 227-247.

Scammacca, N.; Roberts, G.; Vaughn, S.; Edmonds, MS.; Wexler, J.; Reutebuch, CK.; Torgesen, JK. Interventions for adolescent struggling readers: A meta-analysis with implications for practice. RMC Research Corporation; Portsmouth, NH: 2007. CenteronInstruction.Availableathttp://www.centeroninstruction.org/files/COI%20Struggling%20Readers.pdf

Shinn J, Deno SL, Espin CA. Technical adequacy of the maze task for curriculum-based measurement of reading growth. Journal of Special Education. 2000; 34:164–172.

Shinn, MR.; Shinn, MM. AIMSweb training workbook: Administration and scoring of reading maze for use in general outcome measurement. Edformation; Eden Prairie, MN: 2002. Retrieved from http://www.aimsweb.com/uploads/pdfs/scoring_maze.pdf

Texas Education Agency. Appendix 10 – Technical digest 2004–2005. 2004. Retrieved January 8, 2000, from http://www.tea.state.tx.us/index3.aspx?id=4397&menu_id=793

Torgesen, JK.; Myers, D.; Schirm, A.; Stuart, E.; Vartivarian, S.; Mansfield, W.; Haan, C. National assessment of the Title I interim report to congress: Volume II: Closing the reading gap, first year findings from a randomized trial of four reading interventions for striving readers. U.S. Department of Education, Institute of Education Sciences; Washington, DC: 2006. Retrieved from http://www2.ed.gov/rschstat/eval/disadv/title1interimreport/vol2.pdf

Torgesen JK, Wagner RK, Rashotte CA. The prevention and remediation of sever reading disabilities: Keeping the end in mind. Scientific Studies of Reading. 1997; 1:217–234.

Torgesen, JK.; Wagner, RK.; Rashotte, CA. Test of word reading efficiency. Pro-Ed; Austin, TX: 1999.

Vadasy PF, Sanders EA, Peyton JA, Jenkins JR. Timing and intensity of tutoring: A closer look at the conditions for effective early literacy tutoring. Learning Disabilities Research and Practice. 2002; 17:227–241.

Vaughn S, Wexler J, Roberts G, Barth AE, Cirino PT, Romain M, Fletcher JM. The effects of tertiary treatments on middle school students with reading disabilities: Individualized versus standardized approaches. Exceptional Children. in press.

Vaughn S, Cirino PT, Wanzek J, Wexler J, Fletcher JM, Denton CD, Francis DJ. Response to intervention for middle school students with reading difficulties: Effects of a primary and secondary intervention. School Psychology Review. 2010; 39:3–21. [PubMed: 21479079]

Vaughn S, Fletcher J. Thoughts on rethinking response to intervention with secondary students. School Psychology Review. 2010; 39:296–299. [PubMed: 21738283]

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.

Vellutino FR, Scanlon DM, Lyon GR. Differentiating between difficult-to-remediate and readily remediated poor readers. Journal of Learning Disabilities. 2000; 33:223–238. [PubMed: 15505962]

Vellutino FR, Scanlon DM, Sipay ER, Small SG, Pratt A, Chen R, Denckla MB. Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. Journal of Educational Psychology. 1996; 88:601–638.

Wagner, RK.; Torgesen, JK.; Rashotte, CA.; Pearson, N. The test of silent reading efficiency and comprehension. Pro-Ed; Austin, TX: 2010.

Wanzek, J. Response to varying amounts of time in reading intervention for students demonstrating insufficient response to intervention. University of Texas at Austin; 2005. Unpublished doctoral dissertation

Wanzek J, Vaughn S. Research-based implications from extensive early reading interventions. School Psychology Review. 2007; 36:541–561.

Wanzek J, Vaughn S. Students demonstrating persistent low response to reading intervention: Three case studies. Learning Disabilities Research and Practice. 2009; 24:151–163.

Wanzek J, Vaughn S, Roberts G, Fletcher JM. Efficacy of a reading intervention for middle school students identified with learning disabilities. Exceptional Children. in press.

Williams JP. Commentary: Four meta-analyses and some general observations. Elementary School Journal. 2001; 101:349–354.

Wilson, B. Wilson reading system. Wilson Language Training Corporation; Willbury, MA: 1996.

Woodcock, RW.; McGrew, K.; Mather, N. Woodcock-Johnson III tests of achievement. Riverside; Itasca, IL: 2001.

Table 1

Student Outcomes for Treatment Participants in Year 3

	Pretest			Posttest			
Measures	***M***	***SD***	***n***	***M***	***SD***	***n***	**Hedge's *g***
Woodcock-Johnson III Letter-Word Identification	89.39	11.6	28	89.36	9.7	28	−0.01
Woodcock-Johnson III Word Attack	83.00	10.3	28	84.39	9.5	28	0.24
AIMSweb Mazes	93.93	10.0	28	89.58	10.7	28	−0.60
Test of Silent Reading Efficiency and Comprehension	71.50	6.5	28	75.86	10.4	28	0.47
Test of Word Reading Efficiency Phonemic Decoding	85.57	14.7	28	87.07	14.7	28	0.14
Test of Word Reading Efficiency Sight Word	89.93	9.8	28	90.93	10.5	28	0.10
Gates-MacGinitie Passage Comprehension	83.37	8.6	28	82.63	6.3	28	−0.09
Woodcock-Johnson III Passage Comprehension	89.64	8.8	28	88.96	9.4	28	−0.13

Note: All measures are standard scores.

Table 2

Student Outcomes for Comparison Participants in Year 3

	Pretest			Posttest			
Measures	***M***	***SD***	***n***	***M***	***SD***	***n***	**Hedge's *g***
Woodcock-Johnson III Letter-Word Identification	86.62	13.6	13	81.73	12.6	11	−0.59
Woodcock-Johnson III Word Attack	78.31	15.3	13	78.91	11.9	11	−0.08
AIMSweb Mazes	86.41	8.1	13	86.36	5.8	12	−0.02
Test of Silent Reading Efficiency and Comprehension	74.31	11.1	13	72.33	13.8	12	−0.18
Test of Word Reading Efficiency Phonemic Decoding	89.54	15.7	13	80.91	12.6	11	−0.96
Test of Word Reading Efficiency Sight Word	84.08	10.7	13	82.45	9.8	11	−0.07
Gates-MacGinitie Passage Comprehension	85.98	8.4	13	74.48	8.6	12	−0.84
Woodcock-Johnson III Passage Comprehension	88.92	9.9	13	85.36	12.3	11	−0.48

Note: All measures are standard scores.

Table 3

Student Outcomes for Participants Identified as Receiving Special Education in Year 3

	Treatment						Comparison					
	Pretest			Posttest			Pretest			Posttest		
Measures	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***	***M***	***SD***	***n***
Woodcock-Johnson III Letter-Word Identification	79.56	10.2	9	83.00	9.9	9	63.00	1.4	2	60.50	2.1	2
Woodcock-Johnson III Word Attack	84.78	9.2	9	85.89	8.2	9	51.00	4.2	2	57.50	0.7	2
AIMSweb Mazes	90.00	12.3	9	84.55	12.2	9	80.00	7.1	2	85.68	4.8	2
Test of Silent Reading Efficiency and Comprehension	70.11	5.4	9	70.44	8.4	9	65.50	4.9	2	59.00	2.8	2
Test ofWord Reading Efficiency Phonemic Decoding	74.89	12.9	9	77.56	15.7	9	66.50	16.3	*2*	62.00	11.3	*2*
Test of Word Reading Efficiency Sight Word	86.89	11.5	9	85.56	11.0	9	66.0	9.9	2	66.00	2.8	2
Gates-MacGinitie Passage Comprehension	82.83	10.2	9	79.59	7.2	9	74.72	13.6	2	71.51	9.1	2
Woodcock-Johnson III Passage Comprehension	84.56	9.5	9	81.89	9.4	9	70.50	3.5	2	67.00	7.1	2

Note: All measures are standard scores.

Table 4

ANCOVA Results

	Adjusted posttest means				
Year 3 measures	**Tire IV (treatment)**	**Tire I (comparison)**	***F***	***P***	**Hedge's *g***
Woodcock-Johnson III Letter-Word Identification	88.76	83.24	$F(1,36) = 6.564$	.015	0.49
Woodcock-Johnson III Word Attack	83.67	80.76	$F(1,36) = 2.499$	.123	0.24
AIMS web Mazes	87.99	90.07	$F(1,37) = .707$	.406	−0.22
Test of Silent Reading Efficiency and Comprehension	76.48	70.88	$F(1,37) = 2.958$	.094	0.43
Test of Word Reading Efficiency Phonemic Decoding	87.39	80.10	$F(1,36) = 5.114$	.030[a]	0.52
Test of Word Reading Efficiency SightWord	89.68	85.62	$F(1,36)= 1.743$	.195	0.37
Gates-MacGinitie Passage Comprehension	82.78	74.18	$F(1,37)= 12.475$	.001	1.20
Woodcock-Johnson III Passage Comprehension	88.85	85.66	$F(1,36) = 2.209$	.146	0.32

Note: All measures are standard scores.

[a]Not significant with Benjamini-Hochberg Correction of the statistical significance of effects with multiple comparisons.