Hammill Institute on Disabilities

Journal of Learning Disabilities
2017, Vol. 50(2) 195–212

DOI: 10.1177/0022219415618495
journaloflearningdisabilities.sagepub.com
SAGE

The Effects of a Comprehensive Reading Program on Reading Outcomes for Middle School Students With Disabilities

Michael F. Hock, PhD[1], Irma F. Brasseur-Hock, PhD[1], Alyson J. Hock, MS[2], and Brenda Duvel, MA[3]

Abstract

Reading achievement scores for adolescents with disabilities are markedly lower than the scores of adolescents without disabilities. For example, 62% of students with disabilities read *below* the basic level on the NAEP Reading assessment, compared to 19% of their nondisabled peers. This achievement gap has been a continuing challenge for more than 35 years. In this article, we report on the promise of a comprehensive 2-year reading program called Fusion Reading. Fusion Reading is designed to significantly narrow the reading achievement gap of middle school students with reading disabilities. Using a quasi-experimental design with matched groups of middle school students with reading disabilities, statistically significant differences were found between the experimental and comparison conditions on multiple measures of reading achievement with scores favoring the experimental condition. The effect size of the differences were Hedges's g = 1.66 to g = 1.04 on standardized measures of reading achievement.

Keywords

adolescent reading, reading disabilities, reading interventions

The Ongoing Challenge

Overall, the findings reported in the 2013 National Assessment of Educational Progress (NAEP) relative to reading outcomes are encouraging for some learner groups. For example, scores for eighth grade students in general improved slightly. In stark contrast, however, data for eighth grade students with disabilities show that these students continue to do poorly. Fully 62% of students with disabilities read *below* the basic level on the NAEP reading assessment compared to 19% of their nondisabled peers (National Center for Education Statistics, 2013). Students who score at the below basic level are unable to use prior knowledge to make a comparison, describe the central problem faced by a main character, use context to identify meaning of vocabulary, provide text information to support a generalization, read across text to provide an explanation, or support an opinion with text information or related prior knowledge. In a very real sense, students reading at the below basic level are unable to comprehend much of the written material they encounter in school.

The challenges facing students with low reading achievement may become even more pronounced. With the adoption of the Common Core State Standards (CCSS), students with poor reading proficiency will likely face measures with increased rigor (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Thus, students who score poorly on current measures of reading comprehension may score even lower on the more rigorous CCSS aligned measures (e.g., Editorial Board, 2013).

The Need for Solutions

Low levels of reading achievement are also related to poor school outcomes. For example, about 20% of the lowest level readers will drop out of high school by the end of their sophomore year (Dalton, Glennie, Ingels, & Wirt, 2009). Indeed, the number of students with disabilities to graduate from high school actually declined between 1998 and 2005 (Blackorby et al., 2010). The consequences of not graduating have been well documented; dropouts have

[1]Center for Research on Learning, University of Kansas, Lawrence, USA
[2]University of Kentucky, Lexington, USA
[3]Eleanor Roosevelt Middle School, Dubuque Community School District, IA, USA

Corresponding Author:
Michael F. Hock, PhD, Center for Research on Learning, University of Kansas, JRP Hall, 1122 West Campus Rd., Rm. 720A, Lawrence, KS 66045, USA.
Email: mhock@ku.edu

higher unemployment rates and earn lower wages (U.S. Department of Labor, 2003). In sum, academic and social performance gaps for students with disabilities significantly limit educational success and overall life opportunities for this population (Carnegie Council on Advancing Adolescent Literacy, 2010; Lichtenstein & Blackorby, 1995).

The Evidence We Have

Recent studies and meta-analyses of reading interventions and programs for adolescents with reading disabilities have informed our understanding of what works with whom and under what conditions (e.g., Gersten, Fuchs, Williams, & Baker, 2001; Slavin, Cheung, Groff, & Lake, 2008; Swanson, 1999a; Swanson & Hoskyn, 1998; Torgesen, 2005; U.S. Department of Education, Institute of Education Sciences, 2013). Findings from these studies are helping to expand our understanding of the science of reading instruction for adolescents with reading disabilities. We review this research below.

In a meta-analysis of reading research on students with learning disabilities, Swanson (1999b) found that certain principles supporting effective word recognition and reading comprehension had a positive impact on reading outcomes for children and adolescents with learning disabilities. The Swanson study included 92 separate studies conducted between 1972 and 1997. All of the children and adolescents in these studies were identified as having specific learning disabilities. Findings from the analysis indicated that interventions that combined direct instruction and reading strategy instruction were the most effective for teaching reading comprehension. In addition, Swanson found that direct instruction alone was most effective for teaching word recognition. Furthermore, he found that when researchers combined direct instruction and reading strategy instruction, a large effect on student outcomes was obtained (ES = 1.15). Finally, Swanson found that direct instruction of word recognition also had a large effect (ES = 1.06).

In a meta-analysis by Scammacca et al. (2007), researchers studied findings from 31 studies of adolescent struggling readers. Researchers focused their attention on interventions designed to affect reading fluency, vocabulary, and reading comprehension. Researchers sought to answer three primary research questions: (a) How effective were the reading interventions from these studies for adolescent struggling readers? (b) What was the specific impact of these reading interventions on measures of reading comprehension? and (c) What was the specific impact of these reading interventions on students with learning disabilities? Similar to the Swanson analysis, researchers found that teachers could influence reading outcomes for older students with reading difficulties and that adolescents, including those with learning disabilities, could benefit from interventions that target both word-level and reading comprehension strategies. Researchers also found that teaching vocabulary had a positive impact on reading comprehension for older students with reading disabilities.

In a related meta-analysis, Berkeley, Scruggs, and Mastropieri (2010) synthesized findings of 40 studies for improving the reading comprehension of approximately 2,000 students with learning disabilities. Researchers found that the use of reading strategy instruction for students with learning disabilities showed moderate to high effectiveness across the studies. Researchers also found that text enhancements were effective in helping students with learning disabilities learn reading comprehension strategies. These text enhancements included in-text question placement, graphic organizers, and use of technology. Finally, researchers found that teaching students with learning disabilities reading skills using packaged interventions in small groups was an effective method for increasing reading comprehension. These findings supported the analysis conducted by Swanson (1999b) and Scammacca et al. (2007) that reading strategy instruction was effective for students with learning disabilities.

Galuschka, Ise, Krick, and Schulte-Korne (2014) reviewed 22 randomized controlled trials (RCTs) on reading interventions for children and adolescents with reading disabilities. The researchers evaluated 49 comparisons of experimental and comparison groups that included reading fluency, phonemic awareness, reading comprehension, phonics instruction, auditory training, medical treatments, and interventions with colored overlays or lenses. Researchers concluded that phonics instruction was the most frequently used treatment approach that used an RCT design and that phonics instruction was the only approach whose efficacy on reading and spelling performance in children and adolescents with reading disabilities was statistically confirmed. This meta-analysis demonstrated that severe reading and spelling difficulties could be treated with appropriate instructional methods. The authors concluded that systematic instruction of letter–sound correspondences and decoding strategies was the most effective method for improving the literacy skills of children and adolescents with severe reading disabilities.

Sencibaugh (2007) conducted a meta-analysis of 15 studies of reading comprehension interventions for K–12 students with learning disabilities. Results of the synthesis found an effect size of 0.94 for reading comprehension and 1.18 for auditory language strategies. One of the key findings from this analysis was that language dependent strategies had a significant impact on reading comprehension skills of students with learning disabilities. Second, questioning strategies that involved instruction in paragraph restatements along with strategies for text structure analysis yielded the most significant outcomes. The authors also found that the effects of visual and auditory/language strategies produced positive outcomes for students with learning disabilities. Finally, the authors found

that interventions that involved metacognitive strategies produced larger effect size gains. Overall, and consistent with other findings, the results from the synthesis supported the critical importance of teaching students with learning disabilities to effectively use specific strategies to support reading comprehension.

In a synthesis of 29 studies, Gajria, Litendra, Sood, and Sacks (2007) found that several types of interventions were effective in improving the comprehension of expository text by students with learning disabilities (LD). The studies were grouped by intervention type into content enhancement and cognitive strategy instruction. Content enhancement included interventions that used instructional devices (e.g., advance organizers, visual displays, mnemonics, etc.). In all, 11 studies were included in the content enhancement analysis with an overall finding that content enhancement s provided strong support for the use of instructional devices for students with LD. Specifically, researchers found a large effect size ($d = 1.06$) for students with LD on measures of content comprehension.

In the review of cognitive strategy instruction, researchers divided studies into single-strategy and multiple-strategy interventions. Researchers found that the 15 single-strategy interventions from 10 studies resulted in very large gains for students with LD. The effect size gain for single-strategy studies was 1.83. Results for multiple-strategy studies showed even larger gains with an overall effect size of 2.11.

Since the analyses described above were conducted, researchers have continued to study reading instruction for students with learning and reading disabilities. For example, Vaughn, Swanson, and Solis (2013) reported the results of a longitudinal study of reading comprehension interventions for adolescents with LD who were situated within a response to intervention model. In this study, researchers developed interventions across three tiers of instruction, with increasing levels of intensity for students who were nonresponsive to less intense instruction.

The results of the Tier 2 and Tier 3 interventions on student reading achievement scores showed larger gains for the experimental groups than comparison students. However, the impact was considered small ($d = 0.16$). This finding also held for students identified as having an LD, although students with LD performed significantly below the mean of the typical struggling readers in the Tier 3 intervention. This study supported the findings from previous studies but clarified the need for intensive instruction. That is, even with explicit and intensive reading instruction, students with severe reading disabilities demonstrated limited reading improvement.

In another study, Lovett, Lacerenza, De Palma, and Frijters (2012) reviewed the impact of the PHAST PACES reading program with high school struggling readers. Results from this study showed that experimental group students made statistically significant gains when compared to a comparison group. In addition, the overall gains were moderate (Cohen's $d = 0.68$). However, gains deteriorated over time. The authors concluded that while the one semester of PHAST PACES was promising, it did not provide sufficient remediation to narrow the achievement gap for students with severe reading deficits (Lovett, Barron, & Frijters, 2013; Lovett et al., 2012).

In an effort to examine the effects of peer-mediated and teacher-led interventions on high school student reading outcomes, Bemboom and McMaster (2013) used the peer-assisted learning strategies (PALS) model (Fuchs, Fuchs, & Burish, 2010; Fuchs, Fuchs, Mathes, & Simmons, 1997) to determine if a more cost-efficient intervention would be as effective as a teacher-led intervention. In both conditions, students practiced reading fluently and using comprehension strategies. However, in the PALS condition, peers practiced with each other and in the teacher-led condition, the teacher provided multiple models of expert reading and use of reading strategies. Both experimental groups were compared to students in a comparable school. Results of the data analysis showed that there were no statistical differences on the AIMSweb maze measure (Pearson, 2012) between the PALS and teacher-directed groups. In addition, there were no statistically significant differences between the PALS and teacher-directed groups on the AIMSweb oral reading fluency measure. When compared to the nonequivalent comparison group, both experimental groups outperformed the comparison group on the AIMSweb maze test. Specifically, the teacher-directed group had a statistically significant and large effect on the maze text (Cohen's $d = 0.69$) and the PALS group had a statistically significant and very large effect (Cohen's $d = 1.00$) on the same measure.

In a large experimental study designed to measure the effects of integrating teacher knowledge building activities and student-regulated comprehension strategies in 7th to 10th grade English language arts classes, Simmons et al. (2014) reported that integrating teacher knowledge building activities and student-regulated comprehension practices in core English language arts classes was as effective as traditional English language arts instruction for students reading at or below the 30th percentile as measured by the *Gates–MacGinitie* Reading Comprehension subtest. About 5.6% of the students were students with disabilities.

Simmons et al. (2014) found that while there were no statistically significant differences between groups, effect size gains favored the experimental condition. That is, on the fourth edition of the *Gates–MacGinitie Reading Test*, a distal measure of comprehension, researchers found an effect size gain of Hedges's $g = 0.46$. This finding supported previous findings that more intensive instruction may be needed for low performing students (e.g., Vaughn, Roberts, Klinger, Swanson, & Burdman, 2013; Vaughn & Wanzek, 2014; Wanzek, Vaughn, Roberts, & Fletcher, 2011).

The Ingredients for Instructional Success

Taken together, the research described above represents a significant body of evidence-based practices. It is important that the studies and reviews indicate that it is not too late to teach adolescents with disabilities to become better readers and that adolescent struggling readers can benefit in significant ways from reading instruction (Edmonds et al., 2009; Scammacca et al., 2007; Swanson & Hoskyn, 1998; Torgesen, 2005; Vaughn et al., 2013). In addition, this research highlights several common instructional practices that impact reading achievement. These practices include the following: (a) explicit instruction—all studies included some variation of explicit instruction to teach students skills, strategies, or specific knowledge; (b) teaching students multiple cognitive and metacognitive strategies for word-level reading, vocabulary, and reading comprehension; (c) use of cooperative learning activities; (d) teaching reading comprehension in both small groups (about 5 students) and somewhat larger groups (12–15 students); (e) performance feedback by peers or teachers; and (f) standard protocol interventions as individualized interventions for most adolescents with LD. Thus, a foundation of effective instructional practices has been identified and aspects of these effective practices have been integrated into many adolescent reading interventions. However, until recently, few comprehensive interventions have been developed that include all of the effective practices described above.

In 2006, with a grant from the Institute of Education Sciences (PR/Award R305G040011), researchers at the University of Kansas Center for Research on Learning began work to create such an intervention. The resulting program became known as Fusion Reading. Several key features distinguish this program. First, Fusion Reading is comprehensive in scope. It contains instruction that research has identified as necessary for approximately 80% of all struggling readers: phonics, decoding, word recognition, fluency, vocabulary, and comprehension (see Brasseur-Hock, Hock, Kieffer, Biancarosa, & Deshler, 2011; Hock & Brasseur-Hock, 2009; Hock et al., 2009). Second, reading skills and strategies are taught explicitly in a step-by-step, cumulative process. That is, each step of a strategy is taught to mastery, and then all steps are taught together to mastery. This process addresses working memory limitations of many students with reading disabilities. Third, students and their teachers engage in structured cooperative learning activities. While peers practice reading together and take turns as the reader and the coach, the teacher circulates and gives elaborated feedback to students. Thus, peer collaboration is supplemented with teacher instruction and feedback. Fourth, instruction takes place both in small groups and somewhat larger groups. For example, when practicing skills and completing activities, students work either with a partner or in a small group of three to four students. Fifth, Fusion Reading is a standard protocol that can be used with all students, including those with disabilities. Materials are simply adjusted according to individual needs. Finally, teachers provide structured and scaffolded opportunities for students to generalize and apply reading skills and strategies to core class and state-level reading assessment materials. Structured opportunities to apply skills and strategies to actual core materials and assessments and received feedback helps bridge the gap between what happens in a supplemental reading class and the actual reading demands students face in core classes.

Multiple studies have been conducted with Fusion Reading. Among them are the following.

Study 1. As a part of an IES grant (PR/Award R305G040011), an underpowered random assignment study was conducted with Fusion Reading (FR; Hock, Brasseur-Hock, & Deshler, 2012). The control condition was *Second Chance Reading* (SCR; Joyce, Showers, Scanlon, & Schnaubelt, 1998). All students were administered the *Group Reading and Diagnostic Evaluation* (GRADE; Williams, 2001). An independent analysis of the data was conducted. A total of 34 students received instruction in FR, and 35 students received instruction in SCR. The data were analyzed using a hierarchical linear modeling (HLM) approach as implemented in SAS PROC MIXED. The dependent variables were the standard and raw scores on the GRADE comprehension composite test score. A significant interaction was found between treatment and measurement occasion for the standard score on the GRADE Composite comprehension score, $F(2, 88) = 3.53$, $p = .03$. The pre to post gain for the experimental group was statistically significant, $F(2, 88) = 4.59$, $p = .01$. The effect size (Hedges's d) for this subtest score was 0.70, $F(2, 93) = 3.06$, prob = .05, for the raw score and 0.66, $F(2, 93) = 3.73$, prob = .03, for standard scores.

Study 2. A large ($n = 871$ students) RCT was carried out as part of a Striving Readers project. Funded by the Department of Education (DoEd), this study was evaluated independently by SRI (Schiller et al., 2012). The RCT examined the effects of FR, using intent-to-treat, on the reading achievement and motivation of adolescent struggling readers. Eligible struggling readers were assigned randomly to either the FR intervention or a "business as usual" comparison condition. Results were analyzed for only Year 1 of the program; the Striving Readers projects were not funded for Year 2 due to DoEd budget cutbacks. A statistically significant impact was found for the intervention on the Sight Word Efficiency (SWE) subtest of the *Test of Word Reading Efficiency* (TOWRE), with an effect size (Glass Δ) of 0.11. In a supplemental analysis of students who actually received FR, Fusion students had significantly higher TOWRE SWE ($p < .05$, ES = 0.10) and GRADE sentence comprehension

($p < .05$, ES = 0.15) at posttest than comparison group students. There were no significant differences on measures of reading comprehension.

Study 3. A single-group pre-/posttest design study was conducted in a large urban city in the Southeast. All 266 students in the study had individualized education programs (IEPs) and reading goals. In all, 22 middle school teachers participated in the experimental condition. Students were pretested early in the fall of 2012 and post tested 4.5 months later in January 2013. The reading measure was the *Test of Silent Contextualized Reading Fluency* (TOSCRF; Hammill, Wiederholt, & Allen, 2006). Analyses of scores on this measure were calculated for raw score and grade-level scores. The results of the analyses indicated that there were statistically significant differences between pre- and posttest scores for the group ($p = .000$, $d = 0.55$). Grade-level gains as measured by the TOSCRF were statistically significant ($p = .000$, $d = 0.51$), with a mean grade-level gain after 4 months of instruction of close to 1 year (0.92).

Expanding the Evidence

Between 2010 and 2012, another study of FR was conducted to answer the following research questions:

1. What was the impact of the intervention on the reading achievement of middle school students with reading disabilities?
2. What was the magnitude of the gain score difference (effect size) for the intervention when compared to the comparison condition?
3. Was the magnitude of gain scores significantly different for students who received the intervention for 2 years as opposed to students who received the program for 1 year?
4. Did teachers implement the content and instructional practices of the intervention with fidelity?

Two studies evolved from this research. Study 1 took place during the 2010–2011 school year and answered Research Questions 1 and 2. Study 2 took place during the 2011–2012 school year and answered Research Questions 1 and 3. The research surrounding these studies is reported in the following section.

Method

Setting

Both studies were conducted in a medium size urban school district in the Midwest. The district had an enrollment of approximately 10,500 students with two public high schools, three middle schools, and 13 elementary schools. All three middle schools were involved in the studies. The participating middle schools had a student population in which 15.9% of the students had active IEPs, 3% of the students were classified as English language learners, 16% of the students were representative of diverse groups, and 40% of the students were classified as living in low SES households.

Student Participants

A total of 40 sixth grade special education students participated in the studies. They were served in both core classes and in special education supplemental classes. Students were classified by level of special education support provided. For example, students were classified as "more severe" if they were served in special education classes 30% to 50% of their instructional time. Students were classified as "less severe" if they were served in special education classes 10% to 25% of their instructional time. All students had active IEPs that included reading goals. Of the students, 35 (86%) were identified on their IEP as having an LD and 5 (4%) of the students were identified on their IEP as having a hearing impairment. Sign language interpreters were present in all classes that included students with a hearing impairment.

Table 1 presents the sociodemographic characteristics of the study sample. Student participants were recruited from three middle schools. The majority of the students were male ($n = 27$, 67.5%), were Caucasian ($n = 26$, 65.0%), received free or reduced lunches ($n = 22$, 55.0%), and spent 30% to 50% of their time in special education classes ($n = 22$, 55.0%). The sociodemographic characteristics of the sample did not differ between the experimental and comparison groups (see Table 1).

Teacher Participants

A total of 8 sixth grade special education teachers who expressed interest in the studies and signed letters of consent were selected to participate; 5 teachers taught FR, the experimental intervention, and 3 teachers taught Corrective Reading, the district's current reading program and comparison condition. All teachers were certified by the district as highly qualified to teach students with disabilities, had taught reading in the district for at least 2 years, and had taught the district's current reading program during the past school year.

The Experimental Condition

FR was the intervention used in the experimental condition. In this study, FR was taught in small groups of students (3–8) who met regularly in 50-min sessions five times a week throughout the school year.

Table 1. Sociodemographic Characteristics of Study Sample.

	All (N = 40)		Intervention (n = 20)		Comparison (n = 20)					
Variable	*n*	%	*n*	%	*n*	%	χ^2	*df*	*p*	Cramér's *V*
Gender							1.03	1	.311	0.16
Female	13	32.5	8	40.0	5	25.0				
Male	27	67.5	12	60.0	15	75.0				
Race							0.00	1.00	1.000	0.00
Other	14	35.0	7	35.0	7	35.0				
Caucasian	26	65.0	13	65.0	13	65.0				
SES							0.00	1.00	1.000	0.00
Free/reduced lunch	22	55.0	11	55.0	11	55.0				
Regular lunch	18	45.0	9	45.0	9	45.0				
Disability severity							0.40	1.00	.525	0.10
30–50%	22	55.0	10	50.0	12	60.0				
10–25%	18	45.0	10	50.0	8	40.0				

Note. Disability severity is defined as the percentage of time the student was served in special education classes. When not served in special education classes, students were included in general education for the remaining percentage of time.

A 2-year supplemental reading program, FR includes seven instructional units is designed to meet daily for one 50- or 60-min class period each instructional day (Hock et al., 2012). The course does not replace language arts or other core classes but is supplemental to core classes and is usually offered as an elective. The authors suggest that FR be taught in classes consisting of no more than 12 to 15 low-achieving readers in sixth to eighth grade who typically score two or more grade levels below grade placement on a standardized or state reading assessment measure. Four main components are bundled into the program: (a) Word Level Skills, (b) Comprehension, (c) Motivation, and (d) Assessment. Each component is described below.

The Word Level component is taught through The Bridging Strategy (TBS). Bridging consists of four core units: phonics, decoding, word identification, and reading fluency. When students apply TBS, they use multiple skills and strategies to help them quickly and accurately recognize words in text. For example, when students encounter an unfamiliar multisyllabic word, they attempt to say combinations of letters and blend them into a word that is in their listening vocabulary. If that doesn't work, they analyze the beginning and ending letters of the word. Again, students say each word part and blend them together. If the word is still unrecognizable, students proceed to the next step of the strategy and review the remaining letters to find and pronounce the syllable(s). Here, students are taught to find high utility syllable patterns and say each part of the word blending the parts. If students recognize the word, they reread the word in context to check the meaning. If students do not recognize the word, they ask another person, use a dictionary, or use the computer to figure out how to say the word and what the word means. These word skills are taught to a level of automaticity through fluency instruction.

The Comprehension component of FR consists of three key strategies, each of which includes multiple substrategies that support close reading and comprehension. First, students learn the Summarization Strategy. With this strategy, students make multiple passes for multiple purposes through the material by finding clues in reading material, linking the material to prior knowledge, reading short chunks of information, finding main ideas and paraphrasing the ideas, and summarizing sections of text material. Second, students learn the Prediction Strategy. With this strategy, students learn to make meaningful predictions about the text and draw inferences. Finally, students learn the Vocabulary Strategy. This seven-step process includes group, partner, and individual morphological word analysis and extensive discussion and application of context-based vocabulary words. Students also learn how to determine the meaning of unknown vocabulary through the analysis of affixes and context clues and extensive classroom discussion of multiple word meanings, word usage in a variety of contexts, and similarities of the target word to other words.

Two activities embedded in the Comprehension component, Thinking Reading and Book Study, are designed to increase the amount of time disengaged readers spend engaged in the reading process. First, Thinking Reading is an instructional process teachers use to demonstrate expert reading behaviors, to forecast strategy application, and to provide opportunities for students to practice strategy application in the context of authentic reading material. Thinking Reading is similar to Reciprocal Teaching (Palincsar & Brown, 1984) in that the teacher eventually transfers the role of expert reader to students. However, in Thinking Reading, teachers use highly engaging reading materials in an effort to get disengaged readers reengaged with text. Second, Book Study is designed for extension and application of learned

reading strategies outside the classroom. Students select books in their areas of interest and at their independent reading level. Then they complete assignments that are directly related to the strategies and vocabulary being taught. The goals of these activities are to get disengaged readers' "eyes on print" (Vaughn, 2006, p. 172), provide multiple exposures to expert reader models, provide readers with an opportunity to practice new reading strategies, and extend reading practice beyond the classroom.

The Motivation component has several key elements. These elements include the use of highly engaging teen literature, instructional lessons designed for student success, goal setting and performance tracking, and Possible Selves for Readers (PSR). PSR is used to focus students' attention on the importance of becoming an expert reader and, more important, how being an expert reader can help them reach their hopes and dreams as learners, persons, and in a career area. For example, students participate in structured interviews in which they describe themselves as an individual, as a learner, and as a worker. They also identify their hopes, expectations, and fears for the future in each of these areas. These multiple visions of hoped for, expected, and feared selves can be motivating (Markus & Nurius, 1986). From this examination of what is possible for each individual, an action plan is developed that clearly shows the linkage between reading and the attainment of the goals identified by the student. PSR is an ongoing experience and reflects the dynamic nature of student goals. PSR is revisited throughout the 2-year course, with goal progress monitored by both the student and teacher.

The Assessment component is designed to provide individualized data that informs and personalizes instruction. First, formative data are gathered daily for each strategy's various practice activities. For example, checklists are used to measure student use of the target skill or strategy during partner practice activities. A student partner or the teacher watches and listens to the student reader apply a reading strategy to a reading task. The partner or the teacher records whether the student reader used the strategy being taught. This information is used to assess individual student progress and provide immediate, individualized, positive, and corrective feedback to the reader. This information also informs decision making relative to adjusting intervention intensity and difficulty of reading material. These formative data are also used to design modifications to the baseline intervention if students are not making desired progress. Second, progress measures are embedded within each major unit of the curriculum and administered as pre-/postunit tests. These measures inform the learner and teacher as to the level of student mastery of a particular reading strategy, mastery of skills being taught, and comprehension of reading material. Resulting data are used to make program decisions for individual or groups of students. Finally, overall achievement gains are documented by district end-of-grade assessments or other summative standardized reading measures. These data are used to assess the overall impact of FR.

A key structure of FR is the Daily Lesson Format. The Daily Lesson Format provides a structure for the class that ensures all critical instructional activities are included in each class session. For example, during a 60-min class, teachers and students rotate through five activities: Warm-up (5 min), Thinking Reading (12 min), Explicit Instruction (20 min), Vocabulary (18 min), and Wrap-up (5 min). The instructional activities are as follows: (a) Students do a Warm-up activity as soon as they enter the class. The Warm-up is usually a vocabulary question related to the novel the class is reading. Students earn points for completing the activity. (b) Students transition to Thinking Reading, which involves the teacher modeling the behaviors and thinking of an expert reader. Students read highly engaging novels during Thinking Reading and eventually demonstrate and practice reading strategies they are taught in the class. (c) Explicit Strategy Instruction is when teachers explain a strategy, model the strategy, have students practice the strategy, and then provide feedback to students. Students are taught the individual course reading strategies during this time. (d) Next, students study Vocabulary. The process for learning vocabulary is described later. (e) During lesson Wrap-up, students are given a quick assessment of the main skill taught. Usually, this involves having students complete an exit ticket assignment. Also, the next lesson is previewed. The Daily Lesson Format structure helps ensure that each class has instructional variety and that every minute possible is an opportunity for explicit instruction.

A critical and unique feature of the FR is the explicit nature of the generalization and application of reading skills and strategies to core class text material and district assessments. As students learn and practice strategies, they begin with narrative and informational passages written at their instructional reading level. As they become more skillful in using strategies and in understanding the passages that they are reading, they advance to using actual core class materials from language arts, science, social studies, and math. The strategies are used in a flexible manner to respond to unique demands of each core discipline. This final phase of instruction is called Strategy Integration. Strategy Integration occurs three times throughout the course. Thus, students read core materials and apply multiple reading strategies to core materials about 65% of the time, and they also apply the strategies to state assessment practice materials to support generalization of skills and strategies to high stake assessments.

How Professional Development Was Provided

Each FR teacher received extensive professional development (PD) from the program developers. In addition, all building and district-level administrators who were

responsible for curriculum and instruction received PD. The importance of deliberately including building and district leaders in PD plans in urban secondary schools is well documented (e.g., McDonald, Klein, & Riordan, 2009). The model used to provide all PD was based on validated practices for professional learning (e.g., Elmore, 2005; Fullan, 2005; Knight, 2007).

The specific PD provided to FR teachers included the following: (a) Teachers with responsibility for teaching FR were identified as soon as possible and before instruction was delivered. During the fall of Year 1, and prior to the beginning of classes, teachers were provided three days of PD. The PD included information on attributes of struggling readers, theoretical underpinnings of FR, classroom routines and set up, instructional methodology, student grouping strategies, and a review of the instructional materials. In addition, the FR teachers were taught how to instruct students with the first two units of the curriculum: Establish the Course and the Prediction Strategy. (b) During the spring semester of Year 1, FR teachers received three additional days of PD that included instruction on the PSR and TBS. PD included multiple models of experts demonstrating the use of each of the strategies, many opportunities for teachers to practice teaching the strategies and receive feedback, plus opportunities to reflect, ask questions, problem solve, and debrief. (c) During the summer, professional developers were in frequent contact with the curriculum director and special education coordinator to respond to questions, monitor preparation progress, and to provide encouragement and motivation for the upcoming year. (d) In August of Year 2, FR teachers were provided with a full day review of the information from Year 1 and began planning the launch of FR Year 2. (e) Once formal instruction began, extensive instructional coaching was provided to each FR teacher according to the principles of Partnership Instructional Coaching (e.g., Knight, 2007, 2009). (f) During the second year of instruction, teachers had four additional days of face-to-face PD for the purpose of debriefing, problem solving, and learning the remaining components of FR.

The Comparison Condition

The comparison condition for this study was Corrective Reading (CR; see Torgesen et al., 2006). CR is designed to promote decoding, fluency, and comprehension skills of students in Grades 4 to 12 who are reading below their grade level. The program includes four sequential levels that address students' decoding skills and six sequential levels that address students' comprehension skills. The levels are designed to target students who need assistance with particular types of reading skills based on the results of CR placement tests. The decoding and comprehension components can be used separately as a supplemental reading intervention or combined for use as a reading intervention curriculum. All lessons in the program are sequenced and scripted. CR can be implemented in small groups of four to five students or in a whole-class format.

In this study, CR was taught in small groups of students (4–8) that met regularly in 50-min sessions five times a week throughout the school year. Lessons for Year 1 of the studies focused on four sequential levels that addressed students' decoding skills. All teachers in the study who taught CR were experienced in teaching the program and had been provided PD in how to implement the program. Teachers were observed several times throughout the year by their special education department chairperson, who reported that the program was implemented as intended.

Measures

Two measures were used in this study: GRADE (Williams, 2001) and the Northwest Evaluation Association *Measures of Academic Progress* (MAP).

The GRADE is a comprehensive reading diagnostic assessment for grades PK through 12. The Middle Level test (Grades 6–8) includes subtests for Listening Comprehension (LC), Vocabulary (VOC), Sentence Comprehension (SC), and Passage Comprehension (PCO). A composite comprehension (COMP) score is obtained by summing the sentence and passage subtests. In this study, we administered all GRADE subtests except LC. A total test score is obtained by summing the COMP and VOC scores. Vocabulary knowledge is assessed by multiple-choice questions. SC is assessed by choosing a word from a list to fill in a missing word in a sentence. PCO is assessed by reading a short passage and completing multiple-choice questions. The GRADE has high internal consistency reliability estimates ranging from .89 to .99. Alternative form reliability (n = 696) ranged from .81 to .94. Corrected coefficients of Total Reading scores on the *Iowa Test of Basic Skills* and GRADE suggested concurrent validities that ranged between .69 and .83 (n = 185) and from .86 to .90 using the *Gates–MacGinitie Reading Test*. To answer Research Questions 1 and 2, we used the GRADE Total Test score, which is a measure of overall reading achievement.

The MAP tests are computerized adaptive assessments that have been published by the Northwest Evaluation Association (NWEA) since 2000. The purpose of MAP tests is to provide educators with information to inform teaching and learning in reading, mathematics, and science (NWEA, 2011). For this study, we analyzed only the MAP data for the reading portion of the total test score. In general, the MAP reading test consists of 40 multiple-choice items with four options. The MAP reading test comprises four areas, Word Meaning, Literal Comprehension, Interpretive Comprehension, and Evaluative Comprehension, which compose an overall MAP reading

score. For each state, the MAP tests are aligned to specific state content standards by assembling pools of items that address state content standards. The marginal reliabilities of tests across 50 states and grades are consistently in the low to mid-0.90s (NWEA, 2011; Wang, McCall, Jiao, & Harris, 2013). For the analysis in this study, the MAP was used to answer Research Question 3.

Research Design

The research design for this study was a quasi-experimental matched comparison group design with students matched on the following factors: GRADE reading pretest total score and student grade-level placement. Students were also matched on gender, number of hours served in special education classes per day, ethnicity, race, and SES as closely as possible (see Table 1).

Eight special education teachers who expressed interest in the study were assigned by district leadership to teach either FR or the district's current reading program, CR. Five FR teachers each taught one section of FR to a group of three to eight students. All five teachers received the PD described earlier. Three comparison teachers each taught multiple sections of CR to groups of four to eight students. All comparison teachers had previously received PD on how to teach CR.

At the beginning of the school year, 23 students were assigned in small groups to the five FR teachers. The students were asked if they would like to participate in a study of reading programs. All students gave their verbal assent and their parents or guardians subsequently signed consent letters to participate in the study. Before the start of instruction, three students transferred out of the district or had their schedules changed so they were no longer able to participate. Thus, 20 students remained in the FR condition. From a pool of CR comparison students who gave their verbal assent and had their parents or guardians sign consent letters, 20 students were matched on the factors described above with the 20 students in the FR condition. Thus, 40 students were included in the study, 20 in each condition.

Data Analysis

Study 1: GRADE analysis. The GRADE was administered as a pre- and posttest measure during Study 1. Our data had a nested structure such that students were nested within classrooms and classrooms were nested within schools. Thus, we first examined the degree of clustering by calculating intraclass correlations (ICCs) for each of the VOC, SC, PC, and COMP subtest scores of the GRADE at baseline. The average ICC was .00 at the school level (range = 0 to .03), indicating a negligible level of clustering at this level. Clustering at the classroom level was nontrivial with an average ICC of .04 (range = 0 to .15). Single-level analyses were conducted using disaggregation of data at the student level. We confirmed that none of the baseline scores significantly differed either among schools (p = .180 to .906) or among teachers (p = .519 to .950).

Then, repeated-measures analysis of covariance (RM ANCOVA) was performed to examine the changes in the overall GRADE scores over time between students who received the FR intervention and those who received the comparison intervention. The impact of the FR intervention was estimated by a group or group-by-time interaction effect and an effect size was also reported. The control variables included in the model were students' gender, race, SES (free/reduced or regular lunch), disability severity (percentage of time served in special education classes), school, and classroom.

Preliminary assumption testing was conducted to check for normality, linearity, univariate, and multivariate outliers, homogeneity of variance–covariance matrices, and multicollinearity. All analyses were conducted using SPSS version 22.0.

Study 2: MAP analysis. The MAP reading measure was administered across four time points, fall and spring Year 1 and fall and spring Year 2, to all students in the study. Because more than two time points were collected and responses were still nested within students, schools, and teachers, scores were not independent of each other. Multilevel modeling (MLM) was the best tool for data analysis (see Raudenbush & Bryk, 2002). MLM for nested data and similar procedures are common in educational research (e.g., Dettmers, Trautwein, Ludtke, Kunter, & Baumert, 2010; Shen, Leslie, Spybrook, & Ma, 2012). MLM conceptualizes two levels of analysis. The first level (Level 1) consists of each individual's observed development (e.g., the development of MAP scores) over time determined by a set of parameters. The second level (Level 2) consists of individual characteristics that may predict variance in individual's growth over time (e.g., classroom, gender, race, SES, SPED). We used MLM to estimate (a) how MAP scores in this study changed over time, (b) whether individuals demonstrated different trajectories over time and, if they did, (c) whether years of Fusion intervention experience predicted these differences in trajectories, specifically, and (d) whether students who received 2 years of FR had greater gains in MAP scores than students who received only 1 year of the intervention. The comparison group students received CR during Year 1 but received FR instead during Year 2. This was a decision made by the district after an analysis of the Year 1 data.

In the mixed models (models including both Level 1 and Level 2 predictors), we included years of intervention as a

time-varying predictor variable. We included gender, race, SES, SPED, classroom, and school as Level 2 control variables for initial status. Our mixed model was as follows:

$$MAP_{ij} = \pi_{0i} + \pi_{1i}\left(timepoint\right) + \pi_{2i}\left(yearsinterven\right) + \varepsilon_{ij}$$

$$\pi_{0i} = \gamma_{00} + \gamma_{01}\left(race\right) + \gamma_{02}\left(gender\right) + \gamma_{03}\left(SES\right) + \gamma_{04}\left(sped\right) + \gamma_{05}\left(classroom\right) + \gamma_{06}\left(school\right) + \zeta_{0i}$$

$$\pi_{1i} = \gamma_{10} + \zeta_{1i}$$

$$\pi_{2i} = \gamma_{20} + \zeta_{2i}$$

MAP data preparation. Prior to analyses, multivariate assumptions were checked. Next, the prevalence of missing data was examined. Ten subjects were missing data at waves 3 and 4 and three were missing data at all waves. Thus, only 70% of the sample was still participating at the end of the study. Adolescents with missing data did not differ on MAP scores at baseline, $t(35) = 1.25$, $p = .22$. It is therefore not clear whether data are missing completely at random or not missing at random. Missing data were handled with full maximum likelihood. All multilevel analyses were conducted using SAS 9.3.

Descriptive statistics were calculated to describe the sociodemographic characteristics of the study sample (see Table 1). A model was fit predicting MAP scores. All independent variables were categorical and therefore dummy coded. Years of intervention (YRSINTERVEN) was coded as 0 for 0 years, 1 for 1 year, and 2 for 2 years of experience with the Fusion intervention. Gender was a dichotomous variable dummy coded so that 0 referred to the males and 1 referred to females. Race was coded as 0 = Black, 1 = White, 2 = Asian. SES was a dichotomous variable dummy coded so that 0 referred to students who were not receiving free or reduced lunches and 1 referred to students who were receiving free or reduced cost lunches. Severity was a dichotomous variable dummy coded so students who spent 10% to 25% of the time spent in special disability classes were 0 and students who spent 30% to 50% in special disability classes were 1. Effects were considered significant if $p < .05$.

MAP data were fit according to the following model taxonomy: Model A was the unconditional means model and Model B was the unconditional growth model. Next, the Level 2 predictor variables (school, teacher, classroom, students' gender, race, SES (free/reduced or regular lunch and disability severity) were added to the unconditional growth model as control variables. Model D added years of FR intervention as an additional predictor at Level 1 to Model C. Finally, Model E removed the nonsignificant Level 2 predictors from Model D. Delta deviance scores did not show a significant change in model fit when the nonsignificant Level 2 predictors were removed and therefore the more parsimonious model (Model E) was retained.

Fidelity of Implementation

A checklist was developed to measure fidelity of implementation for the experimental condition (see Figure 1). The *Fusion Reading Fidelity Checklist* (FRFC) was divided into two major sections: global fidelity to the lesson format and fidelity to specific instructional procedures. The fidelity measure measured how closely the design of the Daily Lesson Format and instructional practices were followed by FR teachers. The fidelity measure included observation of six essential lesson components: (a) classroom procedures, (b) daily warm-up activity, (c) Thinking Reading activity, (d) direct instruction for reading strategies, (e) vocabulary instruction, and (f) wrap-up and end-of-lesson comprehension check. Teachers were given points for each activity observed for each lesson using a 4-point scoring rubric.

Project researchers trained special education department chairpersons and their designees how to use the FRFC by scoring video of FR teachers and comparing/discussing scores. District staff then observed FR classroom and used the FRFC to measure implementation fidelity. About 15% of the fidelity observations were scored by two observers. Interscorer reliability was determined by the following method: The total number of agreements that the checklist item was observed plus the total number of agreements that the checklist item was not observed were summed. This number was divided by the total number of observable items possible and a percentage of agreement was calculated. The percentage of interscorer reliability was determined to be 86%.

Results

Repeated-Measures Multivariate Analysis of Covariance for the GRADE

Research Questions 1 and 2: Intervention effects for Study 1. Standardized skewness scores and the Shapiro–Wilk test results confirmed normality of data within each group. Also, Levene's test results indicated homogenous variances of each score between the intervention and comparison groups.

An RM ANCOVA was conducted on the overall GRADE scores. There were significant differences between the intervention and comparison group over time, $F(1, 32) = 6.67$, $p = .015$, Hedges's $g = 1.66$ (see Table 2). Pretest and posttest mean and standard deviations for the overall GRADE

Teacher Observed _____Observer(s) _____
Date _____ Unit: _____ Lesson: _____

CLASSROOM CLIMATE **1. Classroom Procedures** **_____ Followed by students?** **_____ Mentioned by teacher or posted?** **_____ Students engaged?** **Time 1 _____** **Time 2 _____** **Time 3 _____**	**COMMENTS:**
LESSON FORMAT **2. Warm-up** **_____ Time allocation followed?** **_____ Warm-up posted?** **_____ Vocabulary-type question?** **_____ Questions about novel?** **_____ Student work reviewed or scored?**	
3. Thinking Reading **_____ Teachers uses strategies while reading aloud?** **_____ Students have "eyes on the page?"** **Time 1 _____** **Time 2 _____** **Time 3 _____** **_____ Students answer questions?**	**NOTE TYPE OF THINKING READING:** **_____ Teacher led? ____ Forecast?** **_____ Teacher guided? ____ Apply?** **_____ Student led? ____ Integrate?**
4. Explicit Instruction **_____ Did the teacher describe a skill or strategy?** **_____ Did the teacher model a skill or strategy?** **_____ Did the teacher provide student practice feedback?**	**NOTE THE LESSON ACTIVITY:** **_____ Vocabulary** **_____ Book study** **_____ Classroom Procedures** **_____ Possible Selves** **_____ Reading strategy taught–** **_____**
5. Vocabulary **_____ 7-step vocab process followed?** **_____ Students engage in discussion?** **_____ 1-2 vocab words taught?**	
6. Wrap-up **_____ Current lesson summarized?** **_____ Next lesson previewed?** **_____ Exit ticket activity conducted?**	

Figure 1. Fusion Reading Fidelity Checklist.

score and individual subcomponents can be found in Table 3. Figure 2 presents the mean scores at baseline and Year 1 follow-up period for the overall GRADE measure. As observed in this figure, the intervention group showed a considerable increase, while the comparison group showed no or little (if at all) changes over time. This pattern of change was supported by significant group-by-time interaction effects and relatively large effect sizes (see Table 2). In summary, these results suggested that the intervention showed a positive impact on students' overall GRADE score, even after controlling for various sociodemographic factors.

A second RM ANCOVA was conducted on MAP scores from Year 1. There were significant differences between the

Table 2. Results of Students' GRADE and MAP Scores After Year 1.

	GRADE				MAP			
Effect	*F*	*df*	*p*	η_p^2	*F*	*df*	*p*	η_p^2
Between-group								
Gender	3.73	1, 32	.062	.105	0.114	1, 27	.739	.004
Race	0.763	1, 32	.389	.023	0.759	1, 27	.391	.027
SES	0.147	1, 32	.704	.005	0.001	1, 27	.974	.000
Severity	0.027	1, 32	.870	.001	0.016	1, 27	.901	.001
School	2.48	1, 32	.125	.072	0.738	1, 27	.398	.027
Classroom	5.13	1, 32	.030	.138	0.001	1, 27	.971	.000
Group	9.60	1, 32	.004	.231	0.157	1, 27	.695	.006
Within-group								
Time	2.34	1, 32	.136	.068	1.93	1, 27	.176	.067
Gender × time	2.02	1, 32	.165	.059	0.099	1, 27	.755	.004
Race × time	0.444	1, 32	.510	.014	0.293	1, 27	.593	.011
SES × time	1.92	1, 32	.176	.057	0.197	1, 27	.661	.007
Severity × time	0.773	1, 32	.386	.024	4.15	1, 27	.052	.133
School × time	0.803	1, 32	.377	.024	4.82	1, 27	.037	.151
Classroom × time	1.29	1, 32	.265	.039	5.87	1, 27	.022	.179
Group × time	6.67	1, 32	.015	.173	5.16	1, 27	.031	.160

Note. df = degrees of freedom (numerator, denominator); GRADE = *Group Reading and Diagnostic Evaluation*; MAP = *Measures of Academic Progress*; SES = socioeconomic status.

Table 3. Pretest and Posttest Means and Standard Deviations for GRADE and MAP Measures After Year 2.

Intervention type	Year 1 Pretest *M*	Pretest *SD*	Year 1 Posttest *M*	Posttest *SD*	Hedges's *g*	Year 2 Pretest *M*	Pretest *SD*	Year 2 Posttest *M*	Posttest *SD*	Hedges's *g*
VOC										
Comparison	7.95	2.70	9.15	4.42	0.328					
Fusion	8.45	2.52	13.70	6.01	1.14					
SC										
Comparison	4.40	2.19	5.20	1.85	0.395				.	
Fusion	4.85	2.11	7.80	2.61	1.24					
PC										
Comparison	8.50	3.02	7.80	4.70	0.177					
Fusion	7.65	2.50	12.0	4.44	1.21					
Overall GRADE										
Comparison	20.85	3.70	21.70	7.31	0.146					
Fusion	20.95	3.22	33.60	10.29	1.66					
MAP										
Comparison	185.47	9.02	188.94	9.40	0.377	197.08	8.42	201.17	7.17	0.523
Fusion	186.20	11.12	195.79	6.90	1.04	187.33	11.38	199.33	8.58	1.19

Note. GRADE = *Group Reading and Diagnostic Evaluation*; MAP = *Measures of Academic Progress*; PC = Passage Comprehension; SC = Sentence Comprehension; VOC = Vocabulary. The GRADE was not administered in Year 2, and the comparison group received the FR intervention in Year 2.

experimental and comparison groups over time; $F(1, 27) = 5.16$, $p = .031$, Hedges's $g = 1.04$ (see Table 2). Pretest and posttest mean and standard deviations for MAP scores are presented in Table 3. The FR intervention improved students' MAP reading assessment score, even after controlling for several sociodemographic factors.

MLM Analysis of MAP Scores After Year 2

Research Questions 1 and 3: Intervention effects for Study 2. Table 4 presents the model results for the model taxonomy. MAP scores were grand-mean centered prior to MLM analysis. For the final model, Model E, initial MAP scores for students at

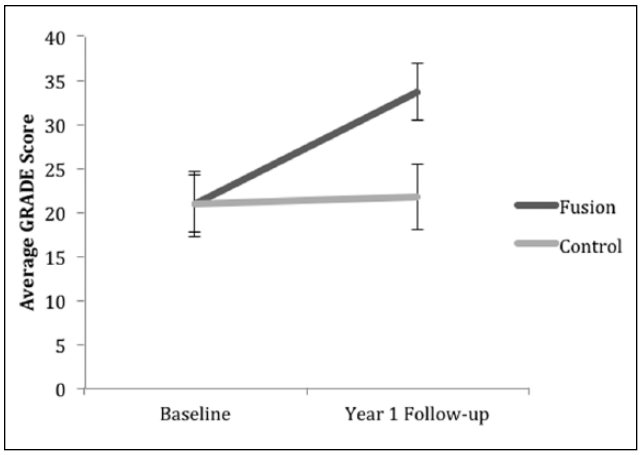

Figure 2. Mean overall Group Reading and Diagnostic Evaluation (GRADE) scores for the intervention and comparison group after Year 1.

"School A" at baseline with 0 years of intervention experience was estimated to be 2.16 points below the mean, on average; however, this initial status was not significantly different from the mean. In addition, the rate of change in MAP scores over time was associated with lower MAP scores, $\pi_{11} = -.439$, $p > .80$, on average, when controlling for school and years of intervention; however, this decrease over time was not significant. Critically, the effect of years of intervention on current MAP scores, controlling for school and changes in MAP over time, was estimated to be 8.00 points higher than baseline scores, $t(26) = 3.53$, $p < .002$. Figure 3 illustrates differences in MAP scores for students with various years of experience with the intervention.

In MLM the effect sizes shown below are generally accepted indices (Raudenbush & Bryk, 2002; Singer & Willett, 2003) but are not comparable in the same sense as a d or η^2. In general, effect sizes tend to fall into two categories: global and local. Global effect sizes quantify the variance in the response variable explained by all predictor variables in an analysis model, whereas local effect sizes quantify the effect of individual variables on the response variable. One way to compute the global effect size statistic R^2 in multiple regression is to compute the predicted score for each participant in the sample, obtain the correlation between the observed and predicted scores, and square that correlation. Although MLM response variable variance is partitioned into Level 1 and Level 2 components, a pseudo-R^2 (see, e.g., Singer & Willett, 2003) for a global effect size statistic can be computed in a similar manner.

Returning to the final model for the current study, the correlation between the observed and the predicted MAP scores was $r = .313$, and squaring this value suggests that 9.80% of the variation in MAP scores can be explained by years of intervention. Below, we present several pseudo-R^2 statistics that quantify how much outcome variation is explained by a multilevel model's predictors. By including years of intervention as a predictor of MAP scores, we have explained 9.80% of the variance in MAP scores from wave 0 to wave 4. In addition, the Level 1 residual variance decreased from 50.19 to 28.81, a difference of 21.38, meaning we have accounted for 42.60% of the within-person variance by adding years of intervention to the model. This indicates that other variables may account for individual differences in the initial status and rate of change of MAP scores among this group of students. To summarize, after 1 year of the FR intervention, students' MAP scores are estimated to be 8.00 points higher than students with 0 years of experience with the Fusion intervention. Furthermore, MAP scores continue to increase with every additional year of intervention.

Research Question 4. Fidelity was assessed in FR classes using the FRFC (see Figure 1). The FRFC is aligned with the Daily Lesson Format described earlier and FR instructional methods. Observers note whether behaviors listed on the FRFC were observed during FR classes. The behaviors were scored using a rubric of 0 to 3 points. If the behavior was not observed during the class, a score of 0 was recorded. A score of 1 point was awarded if the behavior was observed few times. A score of 2 points was awarded if the behavior was observed some times. A score of 3 points was awarded if the behavior was observed most of the time.

Each teacher was observed 3 to 5 times each semester by the special education department chairperson. In addition, at least once each semester, research staff observed FR teachers with district staff and scores were compared to support the reliability of observations. For the five FR teachers, the overall fidelity score was 2.3 points out of 3.0 possible points (see Table 5). Fidelity scores for classroom procedures, warm-up, thinking reading, and direct instruction were high (2.76 to 2.86). Scores for vocabulary and wrap-up were low (1.10 for vocabulary and 1.05 for wrap-up).

Fidelity of implementation for CR was evaluated by observation from the special education chairperson for each school. The observers reported that the comparison reading program was implemented as intended. However, no formal implementation checklist was used to quantify fidelity.

Discussion

The findings from this study of sixth and seventh grade students with reading disabilities indicated that students who received the FR program performed significantly higher on standardized measures of reading than students receiving a comparison program. Specifically, in Study 1, students in the

Table 4. Model Taxonomy Results for MAP Scores After Year 2.

	Model A	Model B	Model C	Model D	Model E
Intercept					
Intercept (π_{00})	−0.141	−5.30**	−6.71	2.54	−2.16
Race (π_{01})			−0.957	−1.11	
Gender (π_{02})			1.20	−1.27	
SES (π_{03})			1.54	3.03	
Sped (π_{04})			1.08	−0.167	
Classroom (π_{05})			0.352	0.772	
School (π_{06})			−1.13	−3.91*	−1.56
Wave					
Intercept (π_{10})		4.02***	4.03***	−0.31	−0.439
Years of Intervention					
Intercept (π_{20})				8.07**	8.00**
Res. var initial	37.54**	46.84*	50.87*	63.11**	59.90**
Res. var wave		3.89	3.87	63.81**	65.59**
Res. var interven				59.19	66.99
Cov (int, wave)		−5.64	−7.99	−18.22	−13.42
Cov (int, interven)				13.10	10.34
Cov (wave, interven)				−66.49	−70.85
L1 res. var	79.06***	50.19***	50.36***	29.14***	28.81***
Deviance	950.60	914.40	913.00	880.5	884.00
Delta deviance		36.20***	1.4	33.90***	3.5

Note. MAP = *Measures of Academic Progress*. Statistical notation provided in parentheses corresponds to the equations provided in the analysis section.
*$p < .05$. **$p < .01$. ***$p < .001$.

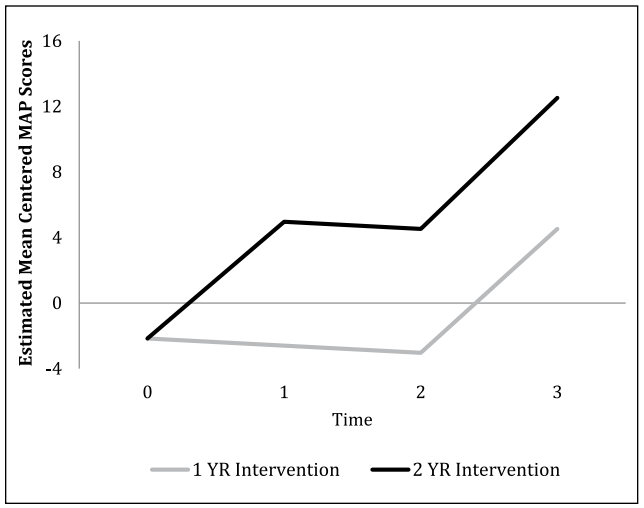

Figure 3. Estimated model-based trajectory of Measures of Academic Progress (MAP) scores at Year 2.

FR program made statistically significant gains for total test scores on the GRADE standardized reading achievement measure and measures of academic progress. The effect size of these gains ranged from Hedges's $g = 1.04$ to 1.66, which are large effects (Cohen, 1988; Richardson, 2011). When one considers that an average effect size gain in reading for students in sixth and seventh grade is 0.23 (Lipsey et al., 2012), the impact of FR on student reading achievement can be substantial.

In Study 2, students who received the FR program for longer periods of time significantly outperformed students who received FR for less time. These results were measured on the MAP reading assessment. Students with 1 year of experience with FR were estimated to score 8.00 points higher on the MAP than students who had not experience with FR. In addition, the MLM model explained 42.60% of the within-person variance. The differences in reading scores for students in Study 2 were statistically significant with 42.60% of the within-person variance of explained by the MLM model. This finding supports conclusions from other studies (e.g., Lovett et al., 2012; Torgesen et al., 2006; Vaughn et al., 2013) that show adolescents with severe reading disabilities may require extended periods of instruction to close reading achievement gaps.

These findings are similar to those of other studies of adolescent reading interventions for middle school students (e. g., Vaughn et al., 2013). However, the magnitude of the gains of students in the FR program shows significant and large effects in overall reading achievement as measured by the GRADE and MAP assessments.

We believe these student gains were the result of several factors. First, due to the structured nature of FR classrooms, teachers were involved in direct instruction activities throughout each lesson and engaged students in learning; there was

Table 5. Fusion Reading Fidelity.

Teacher	Classroom procedures	Warm-up	Thinking reading	Direct instruction	Vocabulary	Wrap-up	Overall fidelity
1	3	2.5	3	3	1	1.3	2.3
2	2.3	2.3	2.3	2.3	1.5	0.5	1.9
3	3	3	3	3	0	2.3	2.4
4	3	3	3	3	3	2.3	2.9
5	3	3	3	3	0	0	2.0
M	2.86	2.76	2.86	2.86	1.10	1.05	2.3

Note. The range of scores possible is 0 to 3 for each category. A score of 0 indicates the behaviors were not observed at all. A score of 1 indicates very few of the behaviors were observed. A score of 2 indicates that some of the behaviors were observed. A score of 3 indicates that most of the behaviors were observed.

no downtime. Second, all of the component reading skills were included in the FR program. That is, students received explicit instruction in phonics, decoding, word identification, fluency, vocabulary, and comprehension over 2 full years. Third, teachers provided scaffolded support as students had multiple opportunities to apply and generalize reading skills and strategies to core class material. Thus, reading strategies were not taught in isolation. Finally, we believe that FR instruction and activities reengaged students with disabilities in the process of reading through the use of culturally responsive, high-interest reading materials and in instruction that was designed for student success.

The data also show the effects of the summer slump or slide. The summer slump or slide is well documented (e.g., Alexander Entwisle, & Olson, 2007; Allington et al., 2010; National Center for Education Statistics, 2011). Specifically, researchers have found that while low-income children made significant gains in reading achievement during the school year, these gains slipped away over the summer due to the lack of summer learning activities (Alexander et al., 2007). Students in both conditions in the current study lost ground over the summer but eventually made up for the loss during Year 2 of the program. For example, Year 1 FR students improved from a fall pretest score of 186.20 on the MAP reading measure to a spring posttest score of 194.55, a gain of 8.35 points. However, when tested with the MAP the following fall, the same students scored 187.33, a loss of 7.22 points. This loss was recovered during Year 2 of FR and students scored 12 points higher on the Year 2 MAP posttest, but may have scored even higher if the summer slide had been addressed.

Limitations

There are several limitations to this study. First, the fact that there is no direct data on fidelity of implementation of the CR program limits the comparison. Whether CR was fully implemented as designed and whether and where instructional overlap occurred between CR and FR is unknown. For example, both CR and FR are explicit instruction models, and explicit instruction has been found to positively impact reading outcomes for students with disabilities (e.g., Swanson, 1999a). Not knowing if CR had been taught explicitly as was designed limits our understanding of what works. Second, the sample size is small and includes only students with IEPs and reading goals, thereby limiting generalizability of the results. In addition, the matched comparison group joined the experimental group, limiting Study 1 to only 1 year of the 2-year program. Thus, the full impact of the 2-year program was not fully tested in Study 1.

Finally, we acknowledge the potential impact of extensive PD and, more important, the potential of instructional coaching to affect student outcomes. Our focus on PD and instructional coaching (Knight, 2007, 2009) may have helped teachers implement FR at a deep level. That is, they came to own FR and made it responsive to the unique context of each school and classroom. They creatively enhanced the program while avoiding lethal mutations to essential components. Thus, we hypothesize that PD is important and instructional coaching is essential for deep implementation and improved student achievement. What level of student outcome growth would have been obtained with traditional PD and less intensive instructional coaching are unknown. The impact of PD and coaching on student reading outcomes is an area in need of future research.

Implications

FR is a supplemental reading program that can be effective if certain systems and structures are in place. For example, a supplemental course requires scheduling support, extensive PD and coaching, a dedicated classroom, and instructional materials. Teachers need extended time to teach; FR requires that students attend the class five times a week for at least 50 min each day. Scheduling challenges in middle and high schools need to be addressed before effective supplemental instruction can be delivered to all students with severe reading disabilities.

We are convinced that there is no short-term solution to the challenge of improving the reading outcomes of students with reading disabilities. Thus, FR is a 2-year curriculum. We think it could be a three-year curriculum covering Grades

6, 7, and 8 to meet the needs of the most severely disabled adolescent readers. Finally, the direct link of supplement reading courses to core class material is critical for generalization of reading skills. Supplemental reading programs that are decontextualized from core class text materials may be one reason for the limited long-term effects of some current reading programs.

The research reported here reflects an effort to incorporate elements of what is known to be effective reading instruction for adolescents with significant reading disabilities into a cohesive and comprehensive reading program. In this program, teachers explicitly teach important word-level skills (i.e., phonics, decoding word recognition, fluency), vocabulary, comprehension, motivational skills, and strategies. In addition, teachers scaffold support as students apply newly acquired reading skills and strategies directly to relevant core class texts and assessment materials. Students actually apply skills and strategies to their core class materials about 65% of the time they are engaged in reading activities.

We feel that this recipe is unique in its explicit instructional methodology, comprehensive scope, motivational component, and structured application of skills and strategies to current core class materials. We believe that this structure and focus account, in large measure, for the promising student learning gains reported in this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Iowa Department of Education (IDOE) and the U.S. Department of Education, Institute of Education Sciences (IES; PR/Award R305G040011). The opinions expressed are those of the authors and do not necessarily represent views of the IDOE or IES.

References

Alexander, K. L., Entwisle, D., & Olson, L. (2007). Lasting consequences of the summer learning gap. *American Sociological Review*, *72*, 167–180.

Allington, R. L., McGill-Franzen, A., Camilli, G., Williams, L., Graff, J., Zeig, J., . . . Nowak, R. (2010). Addressing summer reading setback among economically disadvantaged elementary students. *Reading Psychology*, *31*, 411–427.

Bemboom, C., & McMaster, K. (2013). A comparison of lower- and higher-resourced tier 2 reading interventions for high school sophomores. *Learning Disabilities Research and Practice*, *28*, 184–195.

Berkeley, S., Scruggs, T. E., & Mastropieri, M. A. (2010). Reading comprehension instruction for students with learning disabilities, 1995–2006: A meta-analysis. *Remedial and Special Education*, *32*, 423–436.

Blackorby, J., Schiller, E., Mallik, S., Hebbeler, K., Huang, T., Javitz, H., . . . Williamson, C. (2010). *Patterns in the identification of and outcomes for children and youth with disabilities* (NCEE 2010-4005). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Bloom, H., Hill, C., Black, A. R., & Lipsey, M. (2007, March). *Using empirical benchmarks for interpreting effect size*. Paper presented at the Application of Effect Sizes in Research on Children and Families: Understanding Impacts on Academic Emotional, Behavioral, and Economic Outcomes conference, Washington, DC.

Brasseur-Hock, I. F., Hock, M. F., Kieffer, M. J., Biancarosa, G., & Deshler, D. D. (2011). Adolescent struggling readers in urban schools: Results of a latent class analysis. *Learning and Individual Differences*, *21*, 438–452.

Carnegie Council on Advancing Adolescent Literacy. (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success*. New York, NY: Carnegie Corporation of New York. Retrieved from http://carnegie.org/fileadmin/Media/Publications/PDF/tta_Main.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). New York, NY: Academic Press.

Corrin, W., Somers, M. A., Kemple, J. J., Nelson, E., & Sepanik, S. (2008). *The enhanced reading opportunities study: Findings from the second year of implementation* (NCEE 2009-4036). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Dalton, B., Glennie, E., Ingels, S., & Wirt, J. (2009). *Late high school dropouts: Characteristics, experiences, and changes across cohorts: A descriptive analysis report*. Washington, DC: Institute of Education Sciences.

Dettmers, S., Trautwein, U., Ludtke, O., Kunter, M., & Baumert, J. (2010). Homework works if homework quality is high: Using multilevel modeling to predict the development of achievement in mathematics. *Journal of Educational Psychology*, *102*, 467–482.

Editorial Board. (2013, August 7). New York's Common Core test scores. *New York Times*. Retrieved from http://www.nytimes.com/2013/08/08/opinion/new-yorks-common-core-test-scores.html?_r=0

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, *79*, 262–300. doi:10.3102/0034654308325998

Elmore, R. E. (2005). *School reform from the inside out: Policy, practice, and performance*. Cambridge, MA: Harvard Education Press.

Fuchs, D., Fuchs, L., & Burish, P. (2010). Peer-assisted learning strategies: An evidence-based practice to promote reading achievement. *Learning Disabilities Research and Practice*, *15*, 85–91.

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making class-rooms more responsive to diversity. *American Educational Research Journal*, *34*, 174–206.

Fullan, M. (2005). *Leadership and sustainability: System thinkers in action*. Thousand Oaks, CA: Corwin Press.

Gajria, M., Litendra, A., Sood, S., & Sacks, G. (2007). Improving comprehension of expository text in students with LD: A research synthesis. *Journal of Learning Disabilities*, *40*, 210–225.

Galuschka, K., Ise, E., Krick, K., & Schulte-Korne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: A meta-analysis of randomized controlled trials. *PLOS*, *9*(8), e105843.

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, *71*, 279–320. doi:10.3102/00346543071002279

Hammill, D. D., Wiederholt, J. L., & Allen, E. A. (2006). *Test of silent contextual reading fluency*. Austin, TX: PRO-ED.

Hock, M. F., Brasseur, I. F., Deshler, D. D., Catts, H. W., Marques, J., Mark, C. A., & Wu Stribling, J. (2009). What is the reading component skill profile of adolescent struggling readers in urban schools? *Learning Disability Quarterly*, *32*, 21–38.

Hock, M. F., & Brasseur-Hock, I. F. (2009). Literacy interventions for adolescent struggling readers. In S. R. Paris, D. Fisher, & K. Headley (Eds.), *Adolescent literacy, field tested: Effective solutions for every classroom* (pp. 129–142). Newark, DE: International Reading Association.

Hock, M. F., Brasseur-Hock, I. F., Deshler, D. D. (2012). *Technical report: Fusion reading program*. Lawrence, KS: The University of Kansas Center for Research on Learning.

James-Burdumy, S., Mansfield, W., Deke, J., Carey, N., Lugo-Gil, J., Hershey, A., . . . Faddis, B. (2009). *Effectiveness of selected supplemental reading comprehension interventions: Impacts on a first cohort of fifth-grade students—Executive summary* (NCEE 2009-4058). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Joyce, B., Showers, B., Scanlon, M., & Schnaubelt, C. (1998). A second chance to learn to read. *Educational Leadership*, *55*, 27–30.

Kemple, J., Corrin, W., Nelson, E., Salinger, T., Herrmann, S., & Drummond, K. (2008). *The Enhanced Reading Opportunities Study: Early impact and implementation findings* (NCEE 2008-4015). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Knight, J. (2007). *Coaching: A partnership approach to improving instruction*. Thousand Oaks, CA: Corwin Press.

Knight J. (Ed.). (2009). *Coaching: Approaches and perspectives*. Thousand Oaks, CA: Corwin Press.

Lichtenstein, S., & Blackorby, J. (1995). Who drops out and what happens to them? *Journal for Vocational Special Needs Education*, *18*, 6–11.

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., . . . Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms* (NCSER 2013-3000). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research. Available at http://ies.ed.gov/ncser/

Lovett, M., Barron, R., & Frijters, J. (2013). Word identification difficulties in children and adolescents with reading disabilities: Intervention research findings. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 329–360). New York, NY: Guilford.

Lovett, M., Lacerenza, L., De Palma, M., & Frijters, J. (2012). Evaluating the efficacy of remediation for struggling readers in high school. *Journal of Learning Disabilities*, *45*, 151–169.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954–969.

McDonald, J. P., Klein, E. J., & Riordan, M. (2009). *Going to scale with new school designs: Reinventing high school*. New York, NY: Teachers College Press.

National Assessment Governing Board. (2013). *Reading framework for the 2013 national assessment of educational progress*. Washington, DC: Government Printing Office.

National Center for Education Statistics. (2011). *The condition of education 2011*. Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/programs/coe/overview.asp

National Center for Education Statistics. (2013). *The nation's report card: Reading 2013*. Washington, DC: U.S. Department of Education.

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

Northwest Evaluation Association. (2011, January). *Technical manual for measure of academic progress & measure of academic progress for primary grades*. Portland, OR: Author.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*, 117–175.

Pearson. (2012). *AIMSweb technical manual* (R-CBM and TEL). Eden Prairie, MN: Author.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*, 135–147.

Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007). *Interventions for adolescent struggling readers: A meta-analysis with implications for practice*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Schiller, E., Wei, X., Thayer, S., Blackorby, J., Javitz, H., & Williamson, C. (2012). *A randomized control trial of the impact of the Fusion Reading Intervention on reading achievement and motivation for adolescent struggling readers*. Manuscript submitted for publication.

Sencibaugh, J. (2007). Meta-analysis of reading comprehension interventions for students with learning disabilities: Strategies and implications. *Reading Improvement*, *44*, 6–22.

Shen, J., Leslie, J. M., Spybrook, J. K., & Ma, X. (2012). Are principal background and school processes related to teacher job satisfaction? A multilevel study using Schools and Staffing Survey 2003–04. *American Educational Research Journal*, *49*, 200–230.

Simmons, D., Fogerty, M., Oslund, E. L., Simmons, L., Hairrell, A., Davis, J., . . . Fall, A.-M. (2014). Integrating content knowledge-building and student-regulated comprehension practices in secondary English language arts classes. *Journal of Research on Educational Effectiveness*, *7*, 309–330.

Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly*, *43*, 290–322.

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

Swanson, H. L. (1999b). Reading research for students with LD: A meta-analysis of intervention outcomes. *Journal of Learning Disabilities*, *32*, 504–532.

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, *68*, 277–321. doi:10.3102/00346543068003277

Torgesen, J. K. (2005). *Recommendations for actions to accelerate the reading development of struggling readers in Florida's middle schools*. Tallahassee: Florida Center for Reading Research.

Torgesen, J., Myers, D., Schirm, A., Stuart, E., Vartivarian, S., Mansfield, W., . . . Haan, C. (2006). *National assessment of Title I: Interim report. Volume II: Closing the reading gap: First year findings from a randomized trial of four reading interventions for striving readers*. Washington, DC: National Center for Education Evaluation and Regional Assistance.

U.S. Department of Education, Institute of Education Sciences. (2013). *What Works Clearinghouse: Talent development middle grades program* (Mathematica Policy Research under contract ED-07-CO-0062). Retrieved from http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=617 and http://ies.ed.gov/ncee/wwc/interventionreport.aspx?sid=434

U.S. Department of Labor. (2003). *Statistical fact sheet: Wages and hour fiscal year 2003 enforcement continues record climb*. Retrieved from http://www.dol.gov/whd/statistics/200318.htm

Vaughn, S. (2006). An interview with Sharon Vaughn: The state of reading research and instruction for struggling readers. *Intervention in School and Clinic*, *41*, 169–174.

Vaughn, S., Roberts, G., Klingner, J., Swanson, E., & Bordman, A. (2013). Collaborative strategic reading: Findings from experienced implementers. *Journal of Research on Educational Effectiveness*, *6*, 137–163.

Vaughn, S., Swanson, E., & Solis, M. (2013). Reading comprehension for adolescents with significant reading problems. In H. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 373–387). New York, NY: Guilford.

Vaughn, S., & Wanzek, J. (2014). Intensive interventions in reading for students with reading disabilities: Meaningful Impacts. *Learning Disabilities Research and Practice*, *29*, 46–53.

Vaughn, S., Wexler, J., Roberts, G., Barth, A. A., Cirino, P. T., Romain, M. A., . . . Denton, C. A. (2011). Effects of individualized and standardized interventions on middle school students with reading disabilities. *Exceptional Children*, *77*, 391–407.

Wang, S., McCall, M., Jiao, H., & Harris, G. (2013). Construct validity and measurement invariance of computerized adaptive testing: Application to measures of academic progress using confirmatory factor analysis. *Journal of Educational and Developmental Psychology*, *3*, 88–100.

Wanzek, J., Vaughn, S., Roberts, G., & Fletcher, J. M. (2011). Efficacy of a reading intervention for middle school students with learning disabilities. *Exceptional Children*, *78*, 73–87.

Williams, K. T. (2001). *GRADE: Group Reading Assessment and Diagnostic Evaluation*. Circle Pines, MN: American Guidance Service.