Learning Disabilities Research & Practice, *0*(0), 1–20

DOI: 10.1111/ldrp.12293

Lessons Learned during the Development and Validation of an Intensive Evidence-Based Reading Intervention for Secondary Students

Jean B. Schumaker
Edge Enterprises, Inc.

The purpose of this article is to describe a 40+ year journey during which researchers have endeavored to design and evaluate an integrated, empirically based reading intervention for secondary students and to chronicle the lessons learned as well as suggest future directions for this type of work. Initial development of the intervention began in the 1980s with individual studies, each focused on one part of the reading process (e.g., decoding, vocabulary, comprehension). The quantitative results of these studies showed significant differences between the experimental and the control group or between baseline and after-instruction conditions. In 2004, an opportunity arose to integrate the empirically validated components within a whole reading program and to have that program independently evaluated. Several scaling-up efforts have resulted. The lessons learned and the challenges inherent in scaling up an intervention are shared. Recommendations for future efforts related to putting research into practice are made.

Thousands of students reach the middle and high school grades with reading skill deficits that hinder their ability to respond to the secondary curriculum. In a seminal work in 1980, researchers found that low-achieving seventh graders[1] were reading on average at the fifth-grade level, and seventh graders with learning disabilities (LD) were reading on average at the fourth-grade level (Warner et al., 1980). Further, they found that these students do not make gains in reading across the next six years of school. In other words, the gap between their reading level and their grade level keeps increasing across the years, such that in the 12th-grade year, the 2-year gap for 7th-grade students has grown to a 7-year gap for low achievers and an 8-year gap for students with LD. This seminal research has since been supported by other researchers (e.g., Ferrer et al., 2015; Francis et al., 1996; Stanovich, 1986). Indeed, because this reading achievement gap has been so intractable, Baye et al. (2018) stated that the poor reading performance of secondary students is one of the most important problems facing educators.

Recent data support this conclusion. In 2015, for example, the government reported that the National Assessment of Educational Progress (NAEP) test results showed that 24% of eighth graders were reading below the basic level, including 42% of Black students, 34% of Latino students, and 63% of students with disabilities (National Center for Educational Statistics, 2015). A few years later, NAEP results showed that the eighth-grade reading scores had dropped in 31 states across the nation compared to the 2017 results (National Center for Educational Statistics, 2019). These declines were seen in student populations representing all levels of achievement. Indeed, the latest results have led to the conclusion by some that no progress has been made in the last 30 years in improving student reading achievement (Camera, 2019; Kamil et al., 2008).

Students who experience these kinds of reading deficits are clearly at a disadvantage when they are required to meet the demands of the secondary curriculum (Deshler et al., 2006). Research has revealed that textbooks in required subject-area courses may be written at levels as high as the 12th- to 17th-grade levels (Putnam, 1988; Putnam et al., 1992; Schumaker & Deshler, 1984). The Lexile Framework for Reading (2022) specifies that reading skills at the 11th- to 12th-grade level are required for college and career readiness (Stenner et al., 2012). Moreover, state and national standards require that students learn how to read and to react to what they have read (National Governors Association for Best Practices & Council of Chief State School Officers, 2010a, 2010b). As many as 48 states require students to take minimal competency exams; 12 states require students to pass these exams to graduate from high school (Gewertz, 2017).

Unfortunately, for students with reading deficits, reading instruction is not part of the secondary curriculum. At that level, the focus of instruction has shifted from skills instruction to subject-area instruction (Alverman & Moore, 1991; Anderson, 2009; Ludwig, 2019; Vaughn & Fletcher, 2012). Secondary students are required to earn credits in science, math, social studies/history, and English literature courses and to read considerable amounts of content to acquire information and solve problems to meet the demands of those courses (Putnam et al., 1992; Schumaker & Deshler, 1984).

Clearly, a mismatch exists between what students are expected to do and the reading skills they are able to use effectively to respond to those requirements (Alvermann, 2002). School personnel do not currently have a proactive and effective means for addressing this problem (Reynolds, 2020, 2021). Indeed, today's secondary school curricula do not focus on reading skills (Alvermann & Moore, 1991;

Anderson, 2009; Ludwig, 2019); further, secondary teachers typically do not have the skills to teach reading (Hempenstall, 2021). Instead, school personnel often resort to providing "homework assistance" whereby students are helped to complete their classwork or homework instead of teaching them the skills that they need to learn. Research has shown that this kind of assistance does not improve students' reading achievement, nor does it improve students' failing grades in subject-area courses (Hock et al., 2001a, 2001b). Indeed, completing homework is only part of the equation used to calculate course grades; students must be able to read and respond to tests at passing levels to demonstrate their knowledge in subject-area courses. They are rarely able to do so if they have serious reading deficits and if homework assistance is their only intervention (Hughes & Schumaker, 1991; Hughes et al., 1993). Sadly, secondary teachers often resort to having students find facts in textbooks and fill out worksheets. Students rarely are required to do more than read for periods of time lasting more than 15 seconds (Alvermann & Moore, 1991). Nevertheless, explicit reading instruction is prescribed for struggling secondary students (Anderson, 2009). Naturally, if students do not receive this instruction, continue to fail their courses, and do not earn the necessary credits toward graduation, they are likely to experience frustration from always facing academic situations in which they feel totally overwhelmed and destined to fail (Biancarosa & Snow, 2004). As a result, these students may begin to act out socially, bring delinquency problems to the school, and eventually drop out (Kamil, 2003).

Toward a Solution

Given that large proportions of the secondary population are not able to read above the basic level, a solution to this national catastrophe must be found. Such a solution must quickly and intensively provide students the reading skills they need in a triage fashion such that they can "get on" with their secondary educations without losing much time. An opportunity arose for the development of just such a solution in 2004 when the American Institutes for Research issued a request for proposals (RFP) for the evaluation of supplemental literacy interventions for ninth graders. A team of researchers responded to the RFP by writing a proposal (Deshler et al., 2004) that was funded for several years to develop and evaluate a literacy program associated with the goal of improving the reading performance of secondary students who are two or more years behind their grade level in reading. The team felt that they were particularly suited to create this literacy program because they had conducted research since the early 1980s (e.g., Schumaker & Deshler, 1992; Tralli et al., 1996) to develop and evaluate the effects of a number of academic interventions for struggling secondary students and secondary students with learning disabilities (LD). Several of those interventions had focused on reading (Fisher et al., 2002; Schumaker et al., 2006).

The RFP required that the proposed year-long literacy program be designed for use with ninth graders who are organized in classes of 12–15 students, with each class taught by one teacher for 225 minutes per week. Further, the program had to be closely tied to the findings and recommendations of reading panels (e.g., the National Reading Panel [2000], the Rand Reading Study Group [Snow, 2002], and Reading Next [Biancarosa & Snow, 2004]). It had to be ready to bring to scale, comprehensive and systematic in design and delivery, and well-grounded in research. Further, the program had to include a number of components, including instruction in phonics, fluency, vocabulary, comprehension, metacognition, intensity of instruction, cooperative learning, and motivation for reading.

THE XTREME READING PROGRAM

The proposed program, which became known as the Xtreme Reading Program (Schumaker et al., 2015), is depicted in Figure 1. At the heart of the program are several types of reading instruction. One of the major components in the reading core (shown in the box on the left side of Figure 1) includes advanced decoding and fluency instruction. This instruction provides the reader advanced word-attack skills and increased reading accuracy, rate, and prosody. This component was deemed vital since previous research (Lenz & Hughes, 1990; Warner et al., 1980; Woodruff et al., 2002) had shown that ninth-grade low achievers and students with LD are reading at the fourth- to fifth-grade level. This means that they have learned basic phonics, decoding skills, and sight words but have not learned how to decode the multisyllablic words that they encounter in their secondary science, social studies, and literature texts, nor have they learned to read with expression. Indeed, instruction in these skills has been highly recommended for struggling adolescent readers (Boardman et al., 2008; Hougen, 2015; Quigley & Coleman, 2020). The #1 recommendation of the Institute for Education Sciences practice guide for teachers of Grades 4 through 9 (Vaughn, 2019) is to provide decoding instruction so that the students will be prepared to read multisyllabic words at the secondary level.

Once students have mastered these skills, they can move on to the other major reading component shown in Figure 1: linguistic comprehension instruction. Linguistic comprehension instruction includes vocabulary instruction and instruction in the strategic processes involved in comprehending a variety of written text structures. As such, it provides the reader with the skills and strategies needed to bring meaning to what is being read and to create new knowledge. As depicted in Figure 1, the interaction of the aligned vocabulary and comprehension instruction creates a synergistic or additive effect that results in learning outcomes greater than those that can be generated by either type of instruction alone. Also, when woven together with decoding skills, these combined reading skills form a strong partnership that results in reading success (Scarborough, 2001).

Theoretical Foundations

The reading components depicted in Figure 1 are supported by two theoretical models of reading. On a global level, the program is based on the Simple View of Reading (Hoover

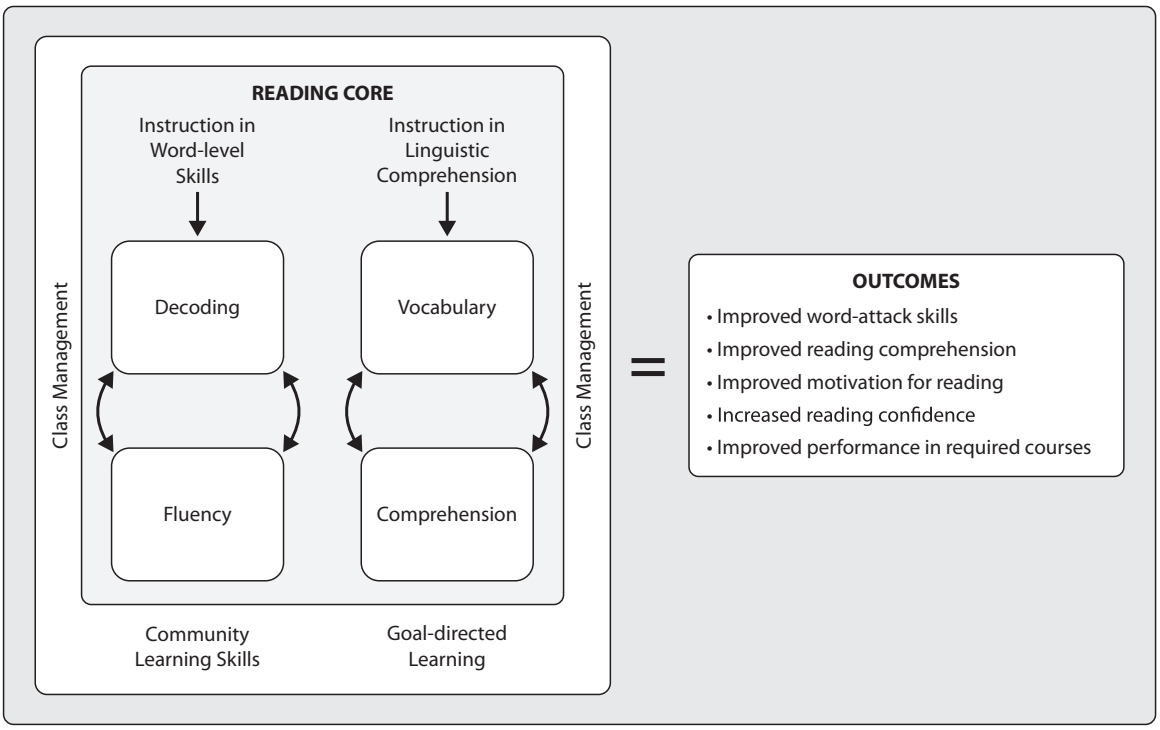

FIGURE 1 The components and potential outcomes of the Xtreme Reading Program.

& Gough, 1990), which holds that, while the act of reading is very complex, proficient reading consists of only two key components. The first component is decoding, and the second component is linguistic comprehension. *Decoding* is defined as efficient word recognition and access to appropriate words in the reader's mental lexicon which provides semantic information at the word level. Efficient decoding allows the reader to quickly pronounce the word and trigger recognition of words acquired through language experiences (e.g., prior knowledge). *Linguistic comprehension*, in turn, is defined as knowledge of facts and concepts, vocabulary, language and text structures, and verbal reasoning structures and strategies. According to the Simple View of Reading, then, the interaction between decoding and linguistic comprehension results in proficient reading comprehension.

On a more specific level, the proposed program components are aligned with a second theoretical model, Kintsch's theory of reading comprehension and learning (Kintsch, 1994, 1998). This theory adds depth to the Simple View of Reading by clearly defining the importance and focus of reading comprehension strategies that enhance the linguistic comprehension component of the Simple View of Reading. The Simple View of Reading lacks clarity with regard to specifying particular strategies that can support proficient reading for those who struggle with reading comprehension. Kintsch (1994) suggested that such strategies can and must be taught to struggling readers, especially when they encounter unfriendly texts (i.e., poorly written or difficult vocabulary) in order to compensate for lack of prior knowledge.

To adhere to Kintsch's theory, the proposed reading program was based on the notion of a "cognitive apprenticeship." (Herein lies the connection of the program to metacognition as required by the original RFP.) An apprenticeship is a relationship between experts and novices in which novices construct knowledge in partnership with experts. While the expert initially leads the instruction, the novice gradually assumes more and more responsibility for applying the knowledge, skills, and strategies introduced by the expert and for choosing the most appropriate strategies to use to fit the circumstances. In the cognitive apprenticeship that comprises the Xtreme Reading Program, the knowledge and cognitive strategies that students (i.e., novice learners) learn are related to the mental processes with which knowledge is acquired. The teacher (the expert learner) scaffolds support while the novice acquires knowledge and skills related to the mental processes involved in literacy (Hock et al., 1993; Pressley et al., 1987b; Pressley & McCormick, 1995; Rogoff, 1990).

The notion of a cognitive apprenticeship has arisen from research that shows that expert learners use cognitive strategies to approach learning tasks. Theorists (e.g., Pressley & McCormick, 1995; Pressley et al., 1987a) have held that these types of strategies should be directly taught to struggling learners to improve their academic performance. Indeed, research over the past 40 years has shown that directly teaching learning strategies to struggling learners does improve their performance on academic tasks (Schumaker & Deshler, 1992; Swanson & Hoskyn, 1998; Swanson et al., 1999). In particular, teaching them reading strategies

improves their performance on reading tasks (Fisher et al., 2002; Fuchs et al., 1999; Shanahan, 2018).

Reading Strategies in the Xtreme Reading Program

The reading strategies that were selected to be taught in the Xtreme Reading Program (Schumaker et al., 2015) align with the reading components shown in Figure 1. Prior to 2004 when the proposal was written, instructional programs for several reading strategies had already been developed and evaluated in separate research studies over 26 years. The Word Identification Strategy, a strategy specifically designed for decoding multisyllabic words (Lenz & Hughes, 1990; Lenz et al., 1984) was selected as a word-level strategy. As needed, prerequisite instruction for this strategy can be individualized to identify those phonic skills needed by individual students, which can be taught to ensure success in mastering the Word Identification Strategy.

In order to improve the students' reading fluency, a repeated-reading process was selected to be used in the Xtreme Reading Program throughout the school year. To implement this process, students repeatedly read instructional-level text to improve their reading accuracy, rate, and prosody for connected text. This repeated-reading process has been developed and empirically validated through a series of research studies (Chard et al., 2002; Fuchs et al., 2001; Kuhn & Stahl, 2003; O'Shea et al., 1985; Torgesen et al., 2001; Wolf & Katzir-Cohen, 2001) as being effective in improving students' reading fluency. Students are taught by direct explanation and expert modeling how to read fluently and how to use the repeated-reading process with partners to read passages aloud and keep track of the number of words read accurately per minute. When students meet or exceed a criterion performance of 140 accurate words per minute plus appropriate prosody in instructional-level text, the text difficulty is gradually increased. They continue this work until they are reading passages in a variety of text types written at their grade level with appropriate rate (i.e., a minimum of 140 words per minute), accuracy (i.e., 98% of the words read accurately), and prosody (i.e., full expression).

With regard to vocabulary instruction in the Xtreme Reading Program, students learn two strategies to enhance their learning of word meanings. The Word Mapping Strategy (Harris et al., 2008, 2011) is a set of cognitive steps students can use to break a word into its most meaningful parts or morphemes, translate those parts into meaning, and then construct meaning from those parts. The second strategy, the LINCS Vocabulary Strategy (Ellis, 1992; Harris et al., 2011), is a set of cognitive steps that enable students to use a group of mnemonic devices to commit the meaning of a word to memory. Toward the beginning of the school year, students learn these strategies and then apply the strategies to 10 new vocabulary words per week. These words are derived from lists of words associated with college entrance examinations as well as words the students may encounter in core secondary courses. Students construct study cards and work in pairs to test each other over the meaning of the words. They take a quiz over the words at the end of each week. Practice and testing are cumulative so that students continue to review previously learned words along with new words each week.

With regard to comprehension, several comprehension strategies are taught within the Xtreme Reading Program. The Visual Imagery Strategy (Clark et al., 1984; Schumaker et al., 1993) is used by students to form mental pictures of events described in a reading passage. The Self-Questioning Strategy (Clark et al., 1984; Schumaker et al., 1994) is used to form questions about information that has not been provided by the author and to find answers to those questions later in the passage. The Paraphrasing Strategy (Ellis & Graves, 1990; Schumaker & Deshler, 1992; Schumaker et al., 1984; Schumaker et al., 2007) is used by students to find the basic structure of a paragraph that contains the main idea and details and to transform the main idea and details into their own words. Finally, the Inference Strategy (Fritschmann et al., 2007a, 2007b) is used by students to make inferences about information in the passage.

All of these comprehension strategies prompt students to activate their prior knowledge and encourage higher order thinking. As each strategy is taught, students also learn about the text structures for which it is most appropriate and how to analyze the components of those text structures, including text structures associated with both narrative and expository text. In addition, they learn how to apply each strategy to high-interest novels, content-area textbooks, and a variety of written materials in their required courses.

The strategies described above are taught within a "spiraling curriculum," meaning that as each new strategy is taught, the instruction for that strategy and the previously taught strategies is integrated. That is, once students learn a new strategy to mastery, they begin using that new strategy integrated with previously learned strategies. They learn how to choose which strategy to use and when.

The Instructional Methodology Used in the Xtreme Reading Program

To ensure that students master learning each strategy, a research-based instructional methodology is used in the Xtreme Reading Program. At the time of the proposal in 2004, this methodology was founded on 26 years of research studies involving 550 research participants, 197 of whom had LD (Schumaker & Deshler, 1992, 2006; Schumaker et al., 2006). The original methodology had eight instructional stages (Ellis et al., 1991) and was designed for instruction with small groups of four to six students. The methodology that evolved for the Xtreme Reading Program to serve 12–15 students per class was based on previous research related to teaching strategies to large classes (Beals, 1985) and consists of eight stages of instruction sandwiched between pretesting and posttesting, or 10 stages in all (see Figure 2).

To begin the instruction, students take a pretest to determine their starting skills. This allows the teacher to quantify each student's performance and facilitates progress monitoring (e.g., Safer & Fleischman, 2005). Next, the teacher describes the steps of the strategy to students and models (or

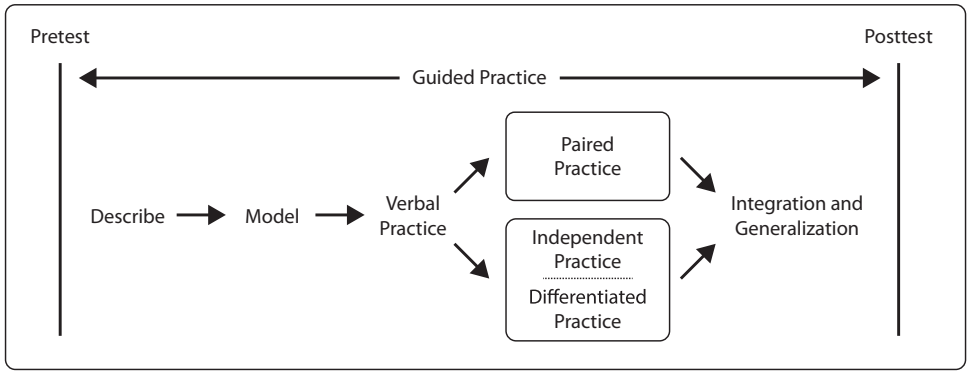

FIGURE 2 The 10 instructional stages for teaching the reading strategies in the Xtreme Reading Program.

demonstrates) using it, thinking aloud so that students can witness all the cognitive processes involved as well as the overt actions. This emphasis on covert behavior helps students to use the covert processes involved in using cognitive strategies (e.g., Roehler & Duffy, 1984) as they imitate the teacher. Subsequently, students learn to say the steps of the strategy from memory through verbal practice. This is important because students may not use self-talk effectively to guide their performances as they use a strategy unless they are taught to do so (e.g., Ellis et al., 1991). They must be able to tell themselves to follow the steps of the strategy in the right order to ensure correct performance.

Once students have mastered naming the steps, they begin to practice using the strategy through three types of practice activities. In paired practice, students practice using the strategy with a partner. In independent practice, students practice by themselves, completing written products demonstrating their use of the strategy. In differentiated practice, students practice applying the strategy aloud with the teacher witnessing all the processes being used and providing feedback to the student. These types of practice are arranged such that students can receive oral or written, positive and corrective feedback from a peer or the teacher each day even though the class is large (12−15 students). This is important because research has shown that certain types of feedback can reduce the number of trials to mastery (Kline et al., 1991). After each practice attempt, students plot their performance on a progress chart. Mastery is required before students progress to learning a new strategy because a strategy must be integrated into students' repertoires at an automatic level (Pressley et al., 1987) if they are to be expected to use the strategy in a generative way.

Once students reach mastery on a given strategy, they begin to practice integrating it with other strategies, and they apply it to their course textbooks and other pertinent materials written at their grade level. This is important because research has shown that students may not generalize their use of a strategy across settings and content areas (Ellis et al., 1989; Schmidt et al., 1988/89). Finally, students take a posttest and celebrate their progress after comparing the posttest results to pretest results.

Throughout these stages of instruction, students engage in daily guided practice, whereby for 10 minutes per class period the teacher leads the whole class while applying a strategy or strategies to the text of a novel. At first, the teacher models how to apply the strategy and then gradually scaffolds student participation in using the strategy (and previously learned strategies) on passages of the novel. This process allows for the daily modeling of covert processes and for the delivery of very timely feedback during an authentic reading task (Islam & Santoso, 2019). Over the course of the school year, six novels are read by the class.

Students also engage in independent book study across the school year where they independently read four books (one per quarter) and complete activities to demonstrate that they have read those books and used the strategies. All of these activities were designed to enable classes of 12–15 students to practice with authentic reading materials and to receive individual feedback on their practice attempts from a peer or the teacher. Such feedback is a critical component of successful strategy instruction (Kline et al., 1991).

The Instructional Environment

In order to ensure that the instruction of reading strategies progresses smoothly in the Xtreme Reading Program, instruction must be surrounded by an environment that promotes and motivates learning (see Figure 1). To create this environment, positive behavioral supports are utilized in the classroom management system to focus student behavior and attention on relevant academic work (Carr et al., 2002; Sugai et al., 2000). In the initial instructional unit, called "Xpect to Achieve," students learn the types of behaviors that will be expected during the various activities that take place in the classroom (e.g., student behaviors that are appropriate during paired practice are not appropriate during independent practice) (Sprick & Garrison, 1998). In addition, students learn community learning skills. First, they learn the SCORE Skills, social skills to be used in cooperative groups or partnerships (Vernon et al., 1993), as well as specific social skills to be used in such activities as whole group discussions (Vernon et al., 2000), partner fluency practice, and partner practice of the reading strategies (Sugai et al., 2000). These skills are critical if students are to be successful in helping each other learn (Johnson & Johnson, 1990;

Johnson et al., 1994). In order to promote students' motivation to learn to read, they also participate in the Possible Selves Program (Hock et al., 2003, 2005) across the school year. Through this goal-directed program, students reflect on their hopes, expectations, and fears, set goals, work toward goals, and attain goals to enhance their future lives. This program is based on the notion that individuals' ideas about what they might become in the future are motivating and can lead to higher academic achievement (Day et al., 1994; Markus & Nurris, 1986).

As depicted on the far right-hand side of Figure 1, the Xtreme Reading Program is intended to result in enhanced outcomes. Students not only learn reading strategies that enable them to decode and comprehend what they read, but their motivation to read increases along with their confidence in themselves regarding being able to respond successfully to the demands of challenging required courses. As they complete those courses successfully, pass minimal competency tests, and graduate from high school, they will be more likely to be able to enroll and succeed in future education and training situations.

Research on Some Individual Components of the Xtreme Reading Program

The original research on four of the reading strategies taught in the Xtreme Reading Program has been selected to illustrate how the instruction for each strategy was validated. The four strategies represent various parts of the instructional sequence across the year-long program: vocabulary building (the Word Mapping and LINCS Vocabulary Learning Strategies), word decoding (the Word Identification Strategy), and comprehension (the Inference Strategy). These strategies have been selected because empirical studies have been published in peer-reviewed journals about the associated research, the studies have been reviewed by the National Center for Intensive Intervention (NCII), and the ratings and reviews are publicly available on the NCII Academic Interventions Tools Chart.[2] The research studies illustrate the kinds of results that can be achieved with instruction in each individual strategy with adolescents.

Vocabulary-Building Strategies

The Strategies

The two vocabulary-building strategies, the Word Mapping Strategy (Harris et al., 2008) and the LINCS Vocabulary Strategy (Ellis, 1992), are strategies that students use to learn the meaning of new vocabulary words. They have different purposes. For example, by using the Word Mapping Strategy, students can predict the meaning of a new word based on the meanings of parts of the word. This strategy was designed based on morphemic analysis instruction (Blachowicz & Fisher, 2000; Spencer, 2000), which involves deriving the meaning of a word by combining the meanings of the parts of the word (Nation, 1990).[3] In contrast, with the LINCS Vocabulary Strategy, students already have been given the meaning of the word (because someone has given it to them, or they have looked it up in a dictionary), and they have been asked to learn/remember that meaning. They do this by using a group of keyword and mnemonic strategies that have been shown to be effective in helping students remember vocabulary (Jitendra et al., 2004). Thus, the Word Mapping Strategy is to be used "on the fly" as students are reading and need to quickly interpret a word, whereas the LINCS Vocabulary Strategy is used deliberately to commit the meaning of a new word to memory using a variety of mnemonic devices or strategies.

To use the Word Mapping Strategy, students first break a word into its morpheme parts (i.e., prefix, suffix, root) and then translate each part into its meaning. Then they stitch together the different meanings of these parts to create a meaning for the whole word. To use the LINCS Vocabulary Strategy, students apply a variety of memory tools to help them remember the meaning of a new word. For example, they think of a "reminding" word that sounds like the new word, they create a story linking the reminding word to the new word, and they draw a picture to link the reminding word to the meaning of the new word.

Both strategies can potentially be used on the same word, but they can be used separately, too. The Word Mapping Strategy is considered to be a generative strategy. That is, by learning this strategy, students can potentially figure out the meaning of thousands of words. Indeed, as many as 50 to 65% of the multisyllablic words that students encounter can be figured out by analyzing their parts (Nation, 1990; Sirles, 1997). In contrast, the LINCS Vocabulary Strategy is more of a mnemonic-based strategy; that is, while using this strategy, each word needs to be learned individually by applying several memory tools (Jitendra et al., 2004).

The Validation Study

Harris et al. (2011) conducted a study on the effects of teaching these two vocabulary strategies to students. The study included 230 ninth graders regularly enrolled in nine English classes. Three classes (the WM Group) ($n = 79$; including 10 SWDs[4]) were taught the Word Mapping Strategy by Harris; three classes (the VL Group) ($n = 79$; including 6 SWDs) were taught the LINCS Vocabulary Strategy, also by Harris; and three classes (the Test-only [TO] Group) ($n = 72$; including 8 SWDs) were taught vocabulary by their own teacher as a part of the regular English curriculum. Within each group, the scores earned by SWDs were examined separately from those of students without disabilities (NSWDs).

When the groups' demographic data were compared using Pearson chi-square analysis, no differences were found with regard to gender ($\chi^2(2) = 3.39$, $p = .184$) and ethnicity ($\chi^2(10) = 4.33$, $p = .931$). Also, no differences among the groups were found with regard to vocabulary standard scores on the SAT-10 (SWDs, $F(2,21) = 2.584$, $p = .099$; NSWDs, $F(2,203) = 2.329$, $p = .100$) and full-scale scores from the WISC-III (SWDs only, $F(2,14) = 1.950$, $p = .179$) when one-way ANOVAs were used.

Several measures were gathered: the Strategy-Use Test, the Word Knowledge Test, and Morphological Analysis Test.

The Strategy-Use Test was used to determine whether the students in the WM Group and the LINCS Group learned the respective strategy they were taught. That is, WM students took a test where they were asked to analyze unknown words and *predict* their meanings; LINCS students took a test where they were asked to *learn* the meanings of unknown words that were supplied to them. The Word Knowledge Test measured students' knowledge of the same 20 words that both the WM and the VL Groups were directly taught during the instruction. They were asked to write any information they connected with the word, to define it, and to use it in a sentence. This test measured their retention of the meaning of a word that had been taught. Finally, the Morphological Analysis Test measured whether students could predict the meaning of words that were supplied to them on the test but not taught during the instruction. This test measured their generative use of the Word Mapping Strategy to unknown words. ANOVAs were used to compare the pretest results for each test across the groups; no differences were found for any test.

Strategy-Use Test. The results on the Strategy-Use Test showed that the students gained skills during the instruction. The students in the WM Group earned significantly higher scores on the posttest than the pretest, Wilks' $\Lambda = .075$, $F(1,77) = 947.03$, $p < .001$, partial $\eta^2 = .925$ (a large effect size). The students in the VL Group also earned significantly higher scores on the posttest than the pretest, Wilks' $\Lambda = .262$, $F(1,77) = 217.184$, $p < .001$, partial $\eta^2 = .738$ (a large effect size). The time by subgroup (SWDs vs. NSWDs) interaction was not significant, meaning that the subgroups of SWDs and NSWDs were gaining at the same rate.

Word Knowledge Test. The results of the Word Knowledge Test showed that both SWDs and nonSWDs in both the WM and LINCS Groups earned posttest scores that were significantly higher than their pretest scores. For SWDs in the WM group, $t(9) = -6.280$, $p < .001$, $d = 4.264$, and for the NSWDs in the WM group, $t(68) = -29.626$, $p < .001$, $d = 8.259$ (both represent very large effect sizes). For the SWDs in the VL group, $t(5) = -5.391$, $p = .003$, $d = 4.226$, and for the NSWDs in the VL group, $t(72) = -26.879$, $p < .001$, $d = 6.299$ (both are very large effect sizes). There were no differences between the posttest scores of the NSWDS in the WM Group and the NSWDs in the VL Group, and also no differences between the posttest scores of the SWDS in the WM Group and the SWDs in the VL Group. These results indicate that each subgroup within the larger intervention groups learned the meaning of the words at similar rates, regardless of their assigned intervention.

Nevertheless, significant differences were found between the Word Knowledge posttest scores of both subgroups in the WM Group and the VL Group and the posttest scores of the respective subgroups in the TO Group. For example, there was a significant difference between the posttest scores of the SWDs in the WM group and the posttest scores of the SWDS in the TO group, $F(1,21) = 24.056$, $p < .001$, partial $\eta^2 = .546$ (a large effect size). Similarly, the posttest scores of the SWDs in the VL group and the posttest scores of the SWDs in the TO group were significantly different, $F(1,20) = 12.589$, $p < .01$, partial $\eta^2 = .386$ (a moderate effect size). These results showed that the WM and VL group students retained the meaning of the same 20 words that they were taught and that the TO group did not learn the meaning of those words.

Morphological Analysis Test. The results of the Morphological Analysis Test showed that the mean posttest scores for the WM students were significantly higher than the posttest scores earned by the VL Group and the TO Group when controlling for the pretest scores. This was true for the SWDs in the WM group vs. the SWDs in the VL group, $F(1,20) = 8.599$, $p < .01$, partial $\eta^2 = .301$ (a large effect size), and the SWDs in the WM Group vs. the SWDs in the TO group, $F(1,20) = 11.801$, $p < .01$, partial $\eta^2 = .371$ (a large effect size). It was also true for the NSWDS in the WM group vs. the NSWDs in the VL group, $F(1,202) = 344.281$, $p < .001$, partial $\eta^2 = .630$, and the NSWDs in the WM Group vs. the NSWDs in the TO group, $F(1,202) = 404.275$, $p < .001$, partial $\eta^2 = .667$ (a large effect size). A significant difference was also found between the posttest and pretest scores of SWDs in the WM Group, $t(9) = -3.45$, $p < .01$, $d = 6.942$ (a large effect size), and for the NSWDs in the WM Group, $t(68) = -21.256$, $p < .001$, $d = 4.646$ (a large effect size). All effect sizes were large. No significant difference between the pretest and posttest was found for either subgroup within the VL and TO Groups. Thus, these results indicate that both subgroups in the WM Group learned how to predict the meaning of significantly more unknown words than the respective subgroups within the VL and TO Groups.

In summary, the Word Mapping Strategy instruction enabled the WM Group to learn to use the Word Mapping Strategy, to retain the meaning of the words they were taught, and to generate accurate meanings for unknown words. The LINCS Vocabulary Strategy instruction enabled the VL Group to learn to use the LINCS Vocabulary Strategy and to retain the meaning of the words that they were taught. "Instruction as usual" with the TO Group did not enable the students to perform at equivalent levels to their peers in the WM Group or the LINCS Group on any test. Both SWDs and NSWDs earned significantly higher scores on the posttests than on the pretests on tests related to the strategy they were taught.

Lessons Learned

The major lessons learned were that, first, if students are to be required to learn and retain the meaning of a word, they need to be directly taught the meaning of the word. This may happen as a result of teaching either the Word Mapping Strategy or the LINCS Vocabulary Strategy at equal levels of effectiveness. The experience of teaching the LINCS Vocabulary Strategy to large classes of students led to the conclusion that large groups of words to which the strategy can be applied needed to be identified and made available to teachers. Use of this strategy is important when students are required to commit to memory the meaning of a list of vocabulary words, a typical task in secondary school (Putnam, 1988). Second, if students are to become adept at predicting the meaning of an unknown word, they need to be taught the Word Mapping Strategy. This skill is important when

students are reading their textbooks and have to quickly decode a word and figure out its meaning. The experience of teaching the Word Mapping Strategy led to the conclusion that lists of words to which the strategy can be easily applied and for which the meaning can be easily predicted needed to be identified so that teachers would not have to search for those words and create lists on their own. Further, for ease in teaching, the words needed to be grouped such that families of words would be grouped according to their suffixes, prefixes, and roots and could be presented together. These conclusions led to the decision to create lists of words that can be grouped together on Learning Sheets that students can work on across the school year. As a result, teachers no longer need to search out the words to teach, and they have lessons that they can teach weekly (see Harris et al., 2008).

The Word Identification Strategy

The Strategy

The Word Identification Strategy (Lenz et al., 1984) was designed to enable students to decode the multisyllabic words that they encounter in the secondary curriculum. It provides students with a set of tools to use when they encounter a long word that they cannot immediately pronounce. That word might be in their speaking or listening vocabulary, but they do not recognize it immediately, or it might be a new word which they have never encountered.

The steps of the Word Identification Strategy enable the student to go through a problem-solving sequence that unlocks the pronunciation of a word. Each step provides a resource known to be helpful in decoding words. In the first step, the student reviews the context within which the word occurs and tries to recall a word that might fit that context and that starts with similar letters (Spear-Swerling & Sternberg, 2002). If that process does not enable the student to pronounce the word, the student identifies and pronounces the prefix and then the suffix. Next, the remaining parts of the word are identified by using a set of rules (called the Rules of Twos and Threes) to divide up the letters and pronounce them. Such an explicit and structured format can be helpful in teaching struggling readers (Henderson & Shores, 1982). At last resort, the student can ask someone for help or check a dictionary.

To teach the Word Identification Strategy, the teacher first introduces the steps of the strategy to students. Next, the teacher models the use of the strategy on a passage, eventually involving the students in the demonstration. Then students learn to name the steps of the strategy from memory. Once they have mastered naming the steps, students practice the steps by applying the steps to reading passages at their reading level. Once they have met a mastery criterion, students begin applying the strategy to more and more difficult passages, increasing the level of the passages gradually until they are applying the strategy at mastery levels at their actual grade level. Finally, students practice applying the strategy in materials being used in their current courses (e.g., textbooks, novels, stories).

The Validation Study

To determine the effects of instruction in the Word Identification Strategy, Lenz and Hughes (1990) conducted a research study with 12 students, aged 13 to 15 years (M = 13.2 years). Four students were seventh graders; two were eighth graders; six were ninth graders. Eight students were White; four were Black. Their IQ scores ranged from 82 to 113. Their reading percentile scores on the *Woodcock Johnson Psychoeducational Battery* (Shrank & Wendling, 2018) ranged from 7 to 32 (M = 15.7). All were diagnosed with LD.

A multiple-baseline-across-students design was employed. Three students participated in each of four multiple-baseline designs; between 11 and 18 data points were gathered across the course of the study for each student for both decoding and comprehension measures. For the decoding data, students orally read a 400-word passage into a recording device. The number of mispronunciations, substitutions, and omissions was counted. Also, the percentage of words correctly pronounced was calculated. (The mastery level was set at 99% correct words.) For the comprehension data, students responded to 10 multiple-choice questions about the 400-word passage they had read the previous day. The percentage of questions answered correctly was determined. Three types of passages were utilized to gather test data from each student during and after the instruction: (a) ability-level passages: passages written at the student's reading level (e.g., if the student's test score showed she was reading at the fourth-grade level, the readability of the passages was at the fourth-grade level); (b) grade-level passages: passages written at the student's grade level (e.g., if the student was in the ninth grade, the readability of the passage was at the ninth-grade level); and (c) textbook passages: passages taken from the textbook used in the student's general-education science class. Thus, students read each type of passage orally and then on the next day answered questions about that passage.

The results showed that the individual students were making decoding errors during baseline. In ability-level materials, six students made fewer than 6 decoding errors, but the remaining students made a mean of 6.3 to 20.5 errors. With regard to the grade-level measure, the whole group made a mean of 12.6 to 37 errors when orally reading 400-word grade-level passages. After instruction on ability-level passages, all the students met mastery (i.e., emitting fewer than 6 errors per passage and reading 99% of the words accurately) within five practice attempts when applying the strategy. The error production decreased as soon as students received instruction in the strategy in each student's case. Four students made no errors. The mean number of errors after instruction ranged from 0 to 6.4 errors in ability-level materials.

Subsequently, all students met mastery within nine practice attempts on grade-level passages. The individual mean number of errors with grade-level materials ranged from 2.9 to 8.3. At the end of the study, the students made a mean of 5 decoding errors as they were reading a 400-word passage taken from their science textbook. Maintenance probes taken as many as five weeks after instruction was terminated

showed that the students maintained their levels of oral reading in grade-level materials at a mean of 5 decoding errors.

With regard to comprehension (i.e., answering comprehension questions correctly), the students' mean comprehension score during baseline was 83% on ability-level materials and 39% on grade-level materials. After instruction in the strategy, the mean comprehension score was 88% on reading-level materials and was 58% on grade-level materials. This means that after decoding instruction, the students were still answering less than 60% of the questions correctly (a failing grade) after they read materials written at their grade levels.

When the Nonoverlap of All Pairs (NAP) statistic (Parker & Vannest, 2009) was used to evaluate the difference between the comprehension data during the baseline and intervention phases of the study (i.e., grade-level practice and maintenance), the percentage of nonoverlap was 75% across all students. In the first multiple-probe design (three students), the percentage of nonoverlap was 84.6% (weighted average Tau = .94). In the second replication (three more students), the percentage of nonoverlap was 76.7 % (weighted average Tau = .89). The percentage of nonoverlap in the third replication was 66.6% (weighted average Tau = .39), and in the fourth replication, the percentage of nonoverlap was 72.5% (weighted average Tau = .66). Effect sizes ranged from Tau = .5 to Tau = 1.0 for individual students.[5] The comprehension gain of nine of the 12 students represented moderate to high effect sizes. Thus, even though the purpose of the study was to test an intervention for reducing the number of word-identification errors made by students, reading comprehension was affected in some students' cases.

The Action-Research Study

An action-research study (Woodruff et al., 2002) was conducted by the staff of a Midwestern high school on the effects of the Word Identification Strategy after the publication of the Lenz and Hughes study (1990). This action-research study is presented here to show how instruction in the strategy might be scaled up within a high school setting for all students who have decoding deficits. Specifically, it shows how the strategy might be taught to larger numbers of students within the same classroom, something that needed to be done to prepare for scaling up the intervention nationally. School personnel in one high school (hereafter referred to as "School A" or the "experimental school") created a decoding program where the Word Identification Strategy was to be taught to any students whose test scores indicated that they were decoding at least two grade levels below their actual grade level. According to the plan, entering ninth graders were to be administered the *Slosson Oral Reading Test—Revised* (Slosson, 1990) at the beginning of school year to determine their decoding skills. Those who met the criterion for inclusion in the program were to be taken out of their English classes daily for four to eight weeks to learn the strategy. Time in the program was to be based on a given student's individual progress to meet the mastery criterion (i.e., fewer than six errors per passage). Groups of students were to rotate into the program as other students who mastered the strategy graduated from the program. Sixty-two ninth graders who met the selection criterion (i.e., decoding at more than 2 years below grade level) were identified in School A.

A matching high school (hereafter referred to as "School B") was identified to serve as the "comparison school." All ninth graders entering School B were administered the *Slosson Oral Reading Test—Revised* (Slosson, 1990) at the beginning of school year to determine their decoding skills. Sixty-two ninth graders in School B who were decoding at least two years below grade level were asked to participate in the study. These students received traditional reading instruction during the school year. Both groups included students who identified as Black, Hispanic, and White as well as students living in poverty. The School A students included 11 students with LD; no students with LD participated in the School B group.

The results of this action-research study replicated the results of the Lenz and Hughes study (1990). That is, once students in School A learned about the strategy and practiced using it, their numbers of oral-reading errors decreased dramatically, and they were able to meet the mastery criterion and graduate from the program. Their improvement was reflected in the standardized decoding test as well. On average, they were decoding at the high fourth-grade level when they entered ninth grade, and they were decoding at the high eighth-grade level at the end of the instruction. In other words, the mean gap between their decoding level and grade level was four to five years at the beginning of the year and less than a year at the end of the school year. They had made an average gain of 3.4 grade levels in decoding skills. The students with LD in the School A Group also improved according to standardized test results. They made an average gain of 3.9 grade levels (range = .8 to 6.1 grade levels). Their average decoding score was at the 4.9 grade level at the beginning of the year and at the 8.8 grade level at the end of the year.

Meanwhile, the students in School B did not make gains in decoding across the school year. At the beginning of the school year, they were decoding at a mean grade level of 6.1. At the end of the year, on average, they were decoding at the 6.3 grade level. The range of mean gain for individual students was .1 to 1.7 grade levels.

ANCOVA results indicated that there was a significant difference between the posttest raw decoding scores of the experimental and comparison groups, $F(1,121) = 31.078$, $p < .001$, $\eta^2 = .692$, a large effect size, when the pretest scores served as the covariate. In addition, there was a significant difference between the posttest scores of students with LD in the experimental group and those of comparison students who were matched with the students with LD in the experimental group on the basis of their pretest scores, $F(1, 19) = 29.673$, MSE = 43.93, $p < .001$, $\eta^2 = .610$. (This also represents a large effect size.) However, the most significant finding was that, at the end of the study, the students in the experimental group were decoding, on average, close to the ninth-grade level, their actual grade level. The experimental students with LD in School A were also decoding at the high eighth-grade level (8.8), very close to their actual

grade level. The results for all subgroups of students, representing all the minorities, mirrored the results for the whole groups. Additionally, girls and boys benefited equally from the strategy program. The authors reported that the school had maintained the decoding program for eight years with similar positive results. Since all the students' comprehension scores did not improve, they also reported that they had begun a similar program for students who needed additional instruction in comprehension.

Clearly, this action-research study has some limitations when compared to the original study (Lenz & Hughes, 1990) that was published in a peer-reviewed journal. First, only two schools were involved, and the teachers chose to do the intervention in their own school with their own students. Although the students were selected in each school according to certain testing criteria, no controls were present regarding who did the testing and how accurate their scoring was. No interscorer reliability was reported, and one might argue that teacher bias was present. Nevertheless, the study illustrates how a decoding intervention can be implemented with a large group of students who enter high school with considerable reading decoding deficits and shows that the school can maintain that intervention across a number of years. The study also informed the developers of the Word Identification Strategy program about further development work that needed to be done to scale up application of the program nationally.

Lessons Learned

Some of the lessons that were learned about teaching the Word Identification Strategy are as follows. First, the teachers who participated in the studies requested graded reading materials in which many multisyllabic words were deliberately embedded in the text to give students more practice using the strategy than possible when using the current teacher-selected materials. They also felt that the students' fluency was poor and requested that fluency exercises be combined with practice of the strategy. Finally, they requested that the materials include both narrative and expository passages to adhere to state reading standards.

As a result, new graded reading passages (for Grades 4 through 10) were written in both narrative and expository genres (Brewer, 2013). Each passage has numerous multisyllabic words embedded in it that can be decoded using the Word Identification Strategy. Also, each passage has a running word count that comprises the right-hand column of each page to facilitate fluency practice. Finally, a fluency exercise was created whereby students work in pairs to read assigned passages aloud to each other while using the strategy, counting the words read within three minutes, and recording their partners' performance on a score sheet. This exercise is used during the instruction of the Word Identification Strategy and weekly thereafter with the goal of students eventually reading at least 140 words per minute in grade-level materials with accuracy and prosody.

However, the biggest takeaway from the Word Identification Strategy studies was that teaching students to decode at a high level of accuracy in grade-level materials does not necessarily improve their reading comprehension. Although some students' comprehension scores improved in the validation study (Lenz & Hughes, 1990), the mean comprehension score for the 12 students was still at the failing level (60%). The researchers decided this outcome was not socially significant even though it was statistically significant. This meant that an effort to develop and validate methods to teach students a variety of comprehension strategies had to be undertaken.

The Inference Strategy

The Strategy

The Inference Strategy (Fritschmann et al., 2007b) is one of the comprehension strategies that was subsequently developed after the Word Identification Strategy as a part of the comprehension package of reading strategies that eventually became part of the Xtreme Reading Program. This strategy was designed to enable students to use a set of comprehension strategies that are likely to contribute to inference generation, including summarization (e.g., Gajria & Salvia, 1992; Malone & Mastropieri, 1992), activating background knowledge and prediction (e.g., Afflerbach, 1990; Billingley & Wildman, 1988), and clarifying (e.g., Simmonds, 1992). Students follow the steps of the Inference Strategy to look for clues in the passage that will inform them about the information that they are seeking and that will help them identify supporting details.

To teach the Inference Strategy, the instructor first introduces the steps of the strategy to students and then demonstrates how to apply the strategy steps to a reading passage and a set of questions to be answered related to the passage. Next, the students engage in activities to learn the names of the strategy steps and information related to those steps and the types of questions they might encounter. Once they have mastered the required information, students practice applying the steps of the strategy to passages written at their reading level and questions associated with those passages. As they become more and more efficient at applying the strategy, they practice with more and more difficult passages until they have mastered applying the strategy to a passage written at their current grade level.

The Validation Study

Fritschmann et al. (2007a) conducted a research study to determine the effects of teaching the Inference Strategy to eight ninth graders, including seven with LD and one with developmental disabilities. Students' ages ranged from 15 years 2 months to 16 years 5 months. Their IQs ranged from 40 to 105. Students' reading scores on the *Group Reading Assessment and Diagnostic Evaluation* (*GRADE*; Williams, 2001) indicated that they were comprehending at least five grade levels below their current grade placement (ranging from the 2.5 to 3.9 grade levels). Seven of them scored at the 1st percentile level, and one scored at the 4th percentile level. Two were White, and the rest represented a variety of ethnicities.

A multiple-baseline-across-students design was used, and the students' performance was measured with several instruments. First, the Strategy Knowledge Test measured whether students retained information about the strategy. This test was comprised of five fill-in-the-blank questions that asked the students to name and explain the steps of the strategy. Students could earn up to 50 points on the test. Additionally, the Strategy-Use Test measured whether the students learned to apply the strategy. To take this test, students read three passages written at their grade level plus the five questions associated with each passage. They earned 1 point each for (a) underlining key words in a question, (b) underlining clue words in the passage related to the question, (c) recording a code letter identifying the *category* of question, and (d) recording a code letter for the *type* of question. Further, they took a criterion-based comprehension test that measured student retention and understanding of the content of each passage. Five multiple-choice questions comprised each test; four of the questions required the student to make a type of inference. Finally, two subtests of a standardized reading test, the *GRADE* (Williams, 2001), were also administered: Sentence Completion and Passage Comprehension. The combined scores from this test represented the comprehension composite score. The measures tracked during the multiple-baseline design consisted of the Strategy-Use Tests and the comprehension tests. The *GRADE* and the Strategy Knowledge Test were administered only twice: at the beginning and end of the study.

On the Strategy Knowledge Test, the students earned a mean score of 0% before instruction and a mean score of 92% after instruction (range = 80% to 100%). Before instruction, the mean percentage score on the Strategy-Use Test was 0%. After instruction, the mean percentage score was 66%, and during the posttest it was 82%. Their scores increased only after they participated in the strategy instruction which was instituted at different times for different students. Significant differences were found between the baseline, posttest, and maintenance conditions using a Friedman Test, $\chi^2(2, N = 6) = 11.565$, $p < .01$. A Wilcoxan signed ranks test (Siegel & Castellan, 1988) was used to conduct pairwise comparisons; and the LSD procedure was used to control Type I errors across these comparisons at the .05 level. The median score for the Strategy-Use posttests was significantly greater than the median score for the baseline tests, $p = .012$.

Regarding the comprehension measure, the mean percentage of comprehension questions the students answered correctly was 32% before instruction. During instruction, the mean percentage score was 77%; during the posttest, the median percentage score was 82%. The students' scores increased only after they participated in the instruction. The median comprehension score earned on the posttests (78%) was significantly higher than the median score earned on the baseline tests (34%), $\chi^2(2, N = 7) = 12.00$, $p < .01$, and the Kendall coefficient of concordance (effect size index) of 1.00 indicated strong differences among the median scores. Follow-up pairwise comparisons revealed a significant difference, $\chi^2(2, N = 7) = 12.00$, $p < .01$, and the Kendall coefficient of concordance (effect size index) of 1.00 indicated strong differences among the median scores. The median score on the comprehension posttests was significantly higher than the median score for the baseline comprehension tests, $p = .012$.

The students' scores on the *GRADE* indicated that they gained an average of 2.8 grade levels in reading comprehension (range = 1.4 to 3.6 grade levels) over the course of the study. Their standard scores on the pretest ranged from 55 to 74 ($M = 61$), and on the posttest from 87 to 94 ($M = 89$). According to the Wilcoxon signed ranks test, a significant difference was found between the posttest and pretest scores, in favor of the posttest scores ($z = -2.521$, $p = .012$), representing a large effect size ($r = 0.91$). At the end of the study, students' grade-equivalent comprehension scores ranged from 4.7 to 7.2 ($M = 5.8$).

Lessons Learned

During the validation study, the need for additional reading matter on which students could practice the strategies became apparent. Not only were graded reading passages corresponding to inferential thinking needed, but passages that were targeted at particular types of inferential thinking were needed. For example, when the students were learning how to determine the main message of a passage, they needed to practice with reading passages that each presented a main message that had to be inferred. A quiz that measured their ability to make that kind of inference was also needed. Thus, reading passages and quizzes were designed to fit each type of inference on which the strategy focuses, so students could practice making that type of inference separately. Additionally, reading passages and quizzes were created for each of grade levels 4 through 10 on which students could practice making a combination of types of inferences (Schumaker & Deshler, 2007).

RESEARCH ON SCALING UP THE XTREME READING PROGRAM

The previous section described some of the research that was conducted on select components of the Xtreme Reading Program before those components had been integrated into a comprehensive adolescent literacy intervention. This section will describe some of the work that was done to prepare the integrated program, as well as the some of the scaling-up research that was conducted after the components were integrated into the Xtreme Reading Program and evaluated as a whole program.

Preparing the Integrated Program

In order to enable teachers to teach the Xtreme Reading Program, instructor's manuals were created for each unit of the program providing step-by-step instructions for the 10 instructional stages to be used in teaching each strategy, with a special emphasis on generalization and integration of the strategies that had been learned. In addition, leveled reading passages were created for each strategy so that students can practice using that strategy and reading for fluency on

more and more difficult passages. A pretest and a couple of posttests were also created for each strategy. Additionally, a Course Organizer (Lenz et al., 1998) was created for the whole program, and a Unit Organizer (Lenz et al., 1994) was created for each instructional unit in the program. These graphic organizers are used to present an overview of the whole course to students and to help them transition between units. Also, a pacing guide was created which includes a monthly plan and daily lesson plans that teachers can follow to stay on pace and complete the whole program within a school year. Finally, a visual depiction of the timeline for the year-long course displaying the strategies to be taught and the novels to be read was also created along with a metaphor for the course that compares participation in the course to participation in an extreme sport. The metaphor is used throughout the course to motivate students to practice hard and perform at their best.

The Original Scaling-Up Study

The first scaling-up study conducted on the Xtreme Reading Program was associated with the grant-funded project mentioned above that was proposed by Deshler et al. (2004). Prior to that project, no evaluation of the Xtreme Reading Program as a whole had been conducted. The project involved 17 school districts in 17 states, spread across the nation. Two high schools participated in each district for a total of 34 schools. The Xtreme Reading Program was taught in one of those high schools, and another reading intervention (the Reading Apprenticeship Academic Literacy [RAAL] Course; Fielding & Schoenbach, 2003) was implemented in the other school. In each school, one English or social studies teacher was assigned to teach the intervention for six class periods per day, five days per week. The teachers were flown to Washington, DC, to participate in week-long professional development (PD) workshops provided by the developers of the reading program to which they were assigned.

At the beginning of the school year, ninth graders were selected by their respective school staffs to participate in the literacy program; they were scheduled to attend the assigned reading course five days per week. The directive to the schools was to choose students whose test scores indicated they were reading at least two or more grades below grade level. Students took the course in place of an elective course. At the beginning and end of the ninth-grade year, all the participating students were bussed to another school and administered the same reading achievement test in the school cafeteria by independent evaluators.

Results were reported by Kemple et al. (2008) and Corrin et al. (2009) for the first and second years of the project, respectively. Implementation issues arose when some teachers failed to start implementing the programs. Only 15 out of 34 teachers started implementing within the first six weeks of school in the first year and earned scores for moderate implementation. The outcomes for students in the 34 schools taken as a whole group ($N = 2916$ for Year 1, and $N = 2679$ for Year 2) showed that the reading comprehension posttest scores were significantly different from the pretest scores in each year (e.g., $p < .05$ for Year 1; $p = 0.042$ for Year 2). The effect size was 0.09 in both years. No significant differences were found between the results of the two interventions. Unfortunately, the test results revealed that 77% of the students were still reading two or more years below grade level at the end of the second project year. In fact, their average reading level was equivalent to a grade 6.1 reading level. Thus, in this case, a statistically significant result did not necessarily represent a socially significant result.

A closer look at the research design, the procedures, and the data analysis reveals some interesting food for thought. First, the research design pitted two reading interventions against each other. There was no "instruction as usual" control within the design or random assignment of students. Second, as reported to the author of the current article (who was directly involved with this study), on testing day, the students were not told where they were going and why. They were simply put on buses and taken to testing sites. Third, implementation of the reading programs was not empirically monitored. According to the professional developers who visited the Xtreme Reading Program sites, some teachers did not implement the program at all. (The professional developers discovered instructor's manuals and student materials stored away in closets in some classrooms months after the school year started.) Some teachers did not follow the pacing guides at all. For example, at Thanksgiving, they were still teaching the vocabulary unit when they should have completed all the work on word-level strategies and moved on to the comprehension strategies. Fourth, many of the students selected for the Xtreme Reading Program were not appropriate for the program. For example, many had scored at the 12th-grade level in reading and had been placed in the course for behavioral issues. Fifth, more than 15 students were assigned to each class period. (The course was designed to be taught to classes of 12–15 students.) Sixth, because the teachers were English and social studies teachers, they were familiar with teaching subject-matter content; they had no experience teaching reading skills and strategies. Some of them had begun to teach literature instead of the reading strategies. For some teachers, the guided-practice periods had devolved to reading a novel aloud to the class. For others, Fridays became silent reading days where students were expected to read silently at their desks, or Fridays were designated as "movie" days.

Finally, because the data were compiled and reported for the whole project, including the students who participated in both interventions, and all the data for all the students regardless of their teachers' implementation were compiled together, a more detailed picture of the results never emerged. Separate from the testing administered by the granting agency, the developers of the Xtreme Reading Program had asked the teachers of that program to administer the *Gates MacGinite Reading Test* (MacGinitie et al., 2000) at the beginning and end of the school year. Thirteen out of 17 teachers did so, and they submitted the test protocols for a total of 502 students. The data were analyzed for three groups of students within this total group. First, the data for the whole group of students were analyzed. Then, because some of the students did not fit the criteria for the program, students whose pretest scores indicated that they

TABLE 1
Mean Raw Scores, Standard Deviations, and Statistics for Responders

Teacher(Students)	*MeanPretest* (SD)	*MeanPosttest* (SD)	F*Value*	p*Value*	*EffectSize* (d)	*Mean Pretest GE*	*Mean Posttest GE*	*MeanGE Gain*
1 ($n = 28$)	25.50 (5.63)	30.71 (6.34)	18.03	<.0001*	.66[b]	6.5	8.2	1.7
2 ($n = 37$)	22.14 (6.001)	27.00 (7.31)	20.74	<.0001*	.62[b]	5.8	7.3	1.5
3 ($n = 27$)	23.96 (6.69)	29.78 (7.67)	21.62	<.0001*	.74[b]	5.7	7.9	2.2
4 ($n = 28$)	20.79 (6.90)	27.00 (6.84)	25.60	<.0001*	.79[b]	5.5	7.3	1.8
5 ($n = 24$)	19.25 (8.23)	29.21 (5.21)	56.36	<.0001*	1.27[c]	5.2	7.9	2.7
6 ($n = 10$)	20.90 (4.20)	25.60 (4.20)	5.23	$p = .023$*	.60[b]	5.6	6.9	1.3
7 ($n = 45$)	20.20 (6.43)	28.18 (6.50)	67.82	<.0001*	1.02[c]	5.4	7.6	2.2
8 ($n = 21$)	22.29 (5.66)	25.81 (7.29)	6.17	.013*	.45[a]	5.9	6.9	1.0
9 ($n = 30$)	22.93 (6.81)	28.27 (7.64)	20.21	<.0001*	.68[b]	6.0	7.6	1.6
10 ($n = 34$)	16.65 (5.49)	28.38 (7.44)	110.88	<.0001*	1.50[c]	4.7	7.6	2.9
11 ($n = 22$)	17.09 (8.997)	27.27 (8.81)	54.01	<.0001*	1.30[c]	4.7	7.4	2.7
12 ($n = 25$)	19.04 (6.71)	28.32 (5.79)	50.98	<.0001*	1.18[c]	5.5	7.6	1.1
13 ($n = 17$)	24.76 (6.74)	29.29 (6.27)	8.26	.004*	.58[b]	6.3	7.9	1.6
Total ($N = 348$)	21.08 (7.01)	28.18 (6.93)	341.06	<.0001*	.91[c]	5.75	7.55	1.8

*Statistically significant.
aSmall effect.
[b]Medium effect.
[c]Large effect.

were reading above the 9.0 grade level were eliminated from the whole group, and the data for the remaining group (those reading at or below grade level on the pretest) (417 students) were analyzed separately. This was done because many of the students who earned scores above grade level were performing at the 11th- and 12th-grade levels. In essence, they had "ceilinged out" on the test, and there was no room to grow. Additionally, they were not the students for whom the Xtreme Reading Program was designed. Next, because the researchers were interested in examining the gains that students who fit the program criteria can make in the program, a third group (called the "Responder Group") was formed of the 348 students who earned scores on the posttest that were at or above their pretest scores. (See Table 1 for this third group's results.)

Clearly, looking at the data in this way after a study has been completed is not conventional, has its limits, and would not be acceptable according to current standards for experimental and quasi-experimental research (e.g., Gersten et al., 2005). Because the data for the Xtreme Reading students were analyzed separately, no control data are available for comparison. Unfortunately, no control group was present in the original study, comparable data were not collected for the RAAL course, and no comparable data are currently available for analysis as a control.

Nevertheless, proceeding with the three groups described above, hierarchical linear modeling (HLM) using the generalized linear model procedure SAS PROC MIXED was applied to each of the three groups separately to compare the pretest to posttest scores. The test scores (pre- and post-) were nested within students, and students were nested within teachers. The results were similar for the three groups. The best fitting model for each of the three groups included time, teacher, and the interaction between teacher and time. All effects were significant for each of the three groups. The posttest scores were significantly higher than the pretest scores; significant differences were found between the mean posttest scores of the different teachers; and the changes that occurred in scores over time were significantly different across the different teachers. For the whole group, follow-up tests using the LSMEANS procedure in PROC MIXED showed that the change from pretest to posttest was significant at the .05 level for nine of the 13 teachers. According to the Responder Group data (see Table 1), effect

sizes for individual teachers ranged from $d = .45$ (slightly below medium) to 1.50 (very large). The students of Teacher 10 had the largest mean gains, averaging 2.9 grade levels, but the students of two other teachers had similar mean gains of 2.7 grade levels. Some of Teacher 10's students made gains of five and six grade levels (Schumaker, 2008). The overall mean gain for all the teachers was 1.8 grade levels. Interestingly, the data in Table 1 replicate the findings of previous research (Warner et al., 1980) showing that many students are reaching the ninth grade reading at the fourth- and fifth-grade reading levels. However, they also show that the deficit gap can be narrowed for some students. Future scaling-up studies need to include a comparison group or implement random selection of students, collect detailed implementation data, analyze the results by teacher for students who fit the program selection criteria, and take testing and implementation factors into consideration for all of the compared groups.

The Portland Striving Readers Project

After the initial scaling-up effort described above, the Portland Oregon schools, as part of the Striving Readers initiative, tested the use of Xtreme Reading in four high schools and five middle schools for five years (Faddis et al., 2011). Students within these schools who met certain selection criteria were randomly assigned to the Xtreme Reading Program or a control group. For example, in Year 1, 348 students were assigned to Xtreme Reading, and 427 students were assigned to a control group. Problems with implementation of the program were reported each year. In Year 1, for instance, only one high school and three middle schools involved teachers with a reading endorsement in the program. Only 40% of the middle school teachers and 25% of the high school teachers attended more than 75% of the PD sessions (Kemple et al., 2008). Although that record improved over the years, some teachers did not attend the PD sessions in Year 5. Implementation fidelity was also an issue. In Year 5, no high schools and only one third of the middle schools were rated as having high levels of fidelity. Participants in Years 1 through 4 were 756 experimental-group students and 823 control-group students.

Results showed that the posttest scores of the experimental students on the GRADE were significantly higher than the control students' posttest scores while controlling for the pretest scores ($\beta = 3.34$, $p < .001$). For the whole experimental group, the standardized effect size was .21. For the middle school students, the standardized effect size was .29; for the high schoolers, it was .12. Also, significantly more treatment-group students than control-group students earned scores representing one grade level or more of improvement at the middle school level, χ^2 (1, $n = 944$) = 9.80, $p < .01$. When the *Oregon Assessment of Knowledge and Skills* (OAKS) test scores were similarly analyzed, no effects were found for the high schoolers. However, significant treatment effects were found for the whole group ($\beta = 1.26$, $p = .05$) and the middle schoolers ($\beta = 1.69$, $p < .05$). Differences in the results were found to be related to implementation fidelity but not to attendance at PD sessions and number of years of teacher experience. In Year 5, district personnel implemented the Xtreme Reading Program in 10 ten additional schools. They included all eligible students in the program instead of randomly assigned students. Thus, data analysis was not conducted for Year 5.

A State Project

On the opposite coast, several schools in one state have evaluated the effects of the Xtreme Reading Program over multiple years led by the state department of education (Boudah, 2018). During the 2013–2014 school year, for example, five middle schools and one high school, representing 23 Xtreme Reading classes with 223 students, contributed data to the evaluation. The results showed that students' scores on the posttest of the *GRADE* and on the *Test of Silent Contextual Reading Fluency* (TOSCRF; Hammill et al., 2006) were significantly higher than their scores on the pretests. With regard to the students' performance on the subtests of the *GRADE* (sentence completion, comprehension, vocabulary), the students' posttest scores on the *GRADE* subtests were significantly higher than their pretest scores on the respective subtests. When student scores were analyzed separately for the students with disabilities, the results showed that these students performed significantly better on the GRADE and TOSCRF posttests than on the pretests ($p = .0001$ in both cases).

Also related to schools in the same state, Boudah (2022) recently reported Xtreme Reading results from seven teachers' classes in six middle schools in two districts. All the teachers participated in PD and received ongoing coaching and support throughout the year. The 141 students included 10 students with disabilities. Again, the students' posttest scores on the *GRADE* and on the *TOSCRF* were significantly higher than their pretest scores. This was also true for each *GRADE* subtest. The mean grade-equivalent (GE) gain was 1.0 GE on the *GRADE* and 2.4 GE on the *TOSCRF*. Similar significant results were found for the students with disabilities on the *GRADE*, but not on the *TOSCRF*.[6]

Although these scaling-up studies have shown positive results, they have their limits. The main limitation is that no experimental controls have been applied. No comparison schools were involved in the studies, nor has random assignment of students been used. Additionally, the teachers were responsible for testing the students at the beginning and end of the year, and no implementation data were collected. Nevertheless, these studies are interesting because they are probably typical of the kinds of evaluations a state or a district might implement in order to make decisions about what interventions to use in the schools.

DISCUSSION

In summary, this article chronicles the initial development and validation of some of the components of a comprehensive reading program that represent examples of the kinds of research that has been conducted and which illustrate how research has led to practice. It also summarizes the results

of some of the scaling-up efforts that have been conducted related to the program. Additionally, some of the lessons learned throughout the 40+ year process of building an integrated and comprehensive reading program and the issues related to the research have been discussed.

One factor that characterizes the research that has been reviewed here is that it was conducted in schools under typical school conditions. Certainly, any type of research in the schools today is difficult. Small studies on individual components are difficult to conduct because instructional time is limited, and school personnel do not want to squander it on an unproven intervention. However, once approval has been obtained for a study, the researcher usually has some control over how and when the instruction takes place. Instructional fidelity can be documented while the instruction is provided. Outcome data can be gathered with some sense of surety by having unbiased individuals administering the assessments and by determining the reliability of the scorers.

In contrast, scaling-up research projects are fraught with challenges at every turn (e.g., Dillenbourg, 2017; Glennan et al., 2004a, 2004b; Quint et al., 2005). By design, they include lots of teachers and schools across several years. Based on the observations of the author of the current article while taking part in many such efforts, administrative and institutional complexities abound when undertaking a scaling-up effort. School administrators who originally approve the project may change their minds or even leave the district. Principals may assign teachers to teach the program who do not want to teach it or have no experience related to the intervention or the type of teaching involved. Sometimes, administrators assign a poorly performing teacher to teach the program to minimize that teacher's contact with larger numbers of students. Principals are unwilling to insist that teachers fully participate in the professional development sessions and implement the program as designed. Principals often do not attend PD sessions to learn about the intervention, nor do they conduct classroom observations. Teachers sometimes refuse to be observed in their classrooms and meet with instructional coaches, citing labor-union rules against meeting before and after school hours. They only implement parts of the program, and do not implement those parts with fidelity. Teachers may even interfere with the gathering of outcome data. The wrong students are assigned to the classes and in larger numbers than can be effectively taught. When these challenges have not been properly addressed by evaluators, proper conclusions cannot be drawn from the data.

Nevertheless, despite these challenges, the research results reported in this article and elsewhere yield some insights related to the Xtreme Reading Program. First, empirical research studies conducted on individual elements of the program have been methodologically sound, they have been published in peer-reviewed journals, and they have shown that student reading performance significantly and substantially improves as a result of students learning a single reading strategy. As a point of fact, the reading gap has been narrowed or closed for students taking part in these studies (e.g., Fritschmann et al., 2007a; Lenz & Hughes, 1990). Second, although the results of scaling-up efforts have been mixed (e.g., see Bouley et al., 2015; Somers et al., 2010; Sprague et al., 2012), and the experimental controls have been lacking in some cases, some of the findings have reflected positive gains (Boudah, 2022; Faddis et al., 2011), and they have been reflected in standardized and state competency test scores. Fourth, the results of the program appear to vary considerably across teachers, with some teachers' students making mean gains of as much as 2.9 grade levels per year and others making mean gains of about one grade level; these differences among teachers and these gains have been replicated across districts and states. Such outcomes reflect the efficacy of the intervention in real classrooms by real teachers who have had very little support or supervision.

Certainly, no one would claim that the scaling-up evaluations of the Xtreme Reading Program have been adequate tests of the program's effects, nor would anyone claim that the evaluations are without limitations. Although the Portland project involved random assignment of the students to an experimental and control group within each school, the poor implementation reports and mixed results between the middle and high schoolers are concerning. Whether this difference is related to poorer implementation at the high school level or the age of the students is not known. Since other projects (e.g., Boudah, 2022; Faddis et al., 2011) yielded positive results in high schoolers, poor implementation is the likely culprit. Nevertheless, the studies in which only pretest and posttest scores were reported with no comparison schools included are problematic and cannot be relied upon as the only data supporting the Xtreme Reading program.

This review of the 40+ year process of building a comprehensive reading program for secondary struggling readers yields plenty of areas where researchers and funding agencies can make improvements in future scaling-up projects for any intervention. A number of factors can increase the probability of success. First, feedback from the teachers indicated that targeted materials need to be created that make the instruction very easy for new teachers. Teachers should not be required to find and compile reading materials for their students that match the reading skills being taught. All the materials should be provided along with lesson plans and other guides. Funding agencies need to supply enough funds that incentives can be provided to school personnel to ensure fidelity of implementation across all dimensions of the program. Administrators need to be given incentives to attend PD to learn about the program and how to monitor its implementation. They also need to receive incentives to ensure that they are responsible and accountable for ensuring that teachers attend PD and implement all parts of the program with fidelity. They need to ensure that the teachers and students selected for the program are appropriate for the program and that the correct number of students are assigned to each class. Teachers need to be given detailed and frequent feedback on the quality and quantity of their instruction tied to incentives for implementing the program, staying on pace, covering all components of the program, and realizing student growth and mastery. Fidelity data need to be collected on the quality of teachers' presentation of the lessons, the percentage of lessons implemented, whether mastery is being required, and the accuracy of the scoring and feedback provided by the teachers. Outcome data should be gathered by personnel not associated with the

school or the program developers. Incentives should be given to students to perform at their best to equal or "beat" their pretest scores to eliminate insincere attempts on the posttest. Ways of analyzing the outcome data need to be used that are tied to teacher implementation levels and related to individual student performance by those students for whom the program was designed. Compiling all the student data together regardless of whether a teacher has implemented the program should not be done. Understanding the individual gains students can make is important when trying to close the achievement gap. Additionally, proper controls (e.g., random assignment of students to the intervention) need to be included to ensure that the studies are accepted by the research community. Other factors such as ensuring that teachers have autonomy in choosing to participate in the program might also be important (Deci, 2009). Indeed, some leaders in the scaling-up field have suggested that a new scaling-up model needs to be adopted; such a model would involve an interactive process that takes into consideration all the factors mentioned above as well as the alignment of policies and infrastructure to ensure the system supports the educational innovation (Glennan et al., 2004a, 2004b).

Notably, the Xtreme Reading Program is not the only secondary-level reading program that has been studied. Numerous other studies have now been completed on other reading interventions at the secondary level (see Baye et al., 2018; Slavin et al., 2008, for reviews). Some of them focus on reading skill/strategy instruction (e.g., Schiller et al., 2012; Vaughn & Fletcher, 2012; Vaughn et al., 2013), while others focus on the method of instruction (e.g., technology [e.g., Shannon & Grant, 2015], intensive group instruction [e.g., Lang et al., 2009], or cooperative learning [e.g., Stevens & Durkin, 1992]). Although the effect sizes achieved in some of the studies are comparable to those in the scaling-up studies reviewed here, there is no standard set of parameters on which the interventions and studies can be compared. Even though two studies have achieved an effect size of .09, no one really knows what that means in terms of student gains within individual teachers' classes. Although suggestions have been made about principles to follow when scaling-up interventions (e.g., Baker, 2004; i3 Community, 2017; Quint et al., 2005), whether these principles have been followed is unknown. Furthermore, the actual meaning of significant differences and effect sizes in terms of the quality-of-life factors important in raising a student's reading level is not clear. There is no way of knowing whether the achievement gap has been closed in a meaningful way by these interventions such that students can succeed in high school and beyond. Whether certain levels of statistical significance and certain effect sizes translate into the ability to perform in high school courses is anyone's guess. When the field identifies standard parameters related to implementation, outcome data, and data analysis, the results of these comparisons may become clear.

In sum, the research community needs to learn from the experiences that have been derived from the scaling-up efforts of the past. Teachers should not be expected to teach intensive and explicit programs aimed at helping students close the deficit gap without appropriate professional learning experiences and supervision, including accountability checks. Such supervision needs to ensure that all parts of the program are being implemented in the proper order and at the proper pace at a high level of quality. Evaluators and educators need to stop assuming that implementation is taking place and do more than rating in a global way whether it is taking place. They need to stop assuming that content-area secondary teachers can be expected to teach reading to students with severe reading deficits without considerable buy-in, professional development, detailed feedback, and incentives.

Perhaps scaling-up efforts can take place in phases with only small numbers of teachers and classes being monitored at first until the training and implementation aspects of the project have been ironed out. Perhaps fewer numbers of components are needed to achieve the same results.[7] Perhaps trying to remediate all of a student's deficits in one year is not practical. Perhaps students can be taught in two years and followed into additional grades to determine the effects on their performance and longevity in school. Perhaps teachers who have the most success can be studied to determine what sets them apart from other teachers. Once some of these issues have been addressed, larger numbers of schools can be included in scaling-up efforts. Only then will adequate scaling-up efforts be achieved with a complex, multicomponent intervention aimed at closing the achievement gap.

Acknowledgments

The author wishes to acknowledge the contributions of Drs. Donald Deshler, Janis Bulgren, and Michael Hock to the original development of the Xtreme Reading Program. She also wishes to express gratitude to all the teachers across the nation who have implemented the program over the years and the professional developers who have provided workshops for and coached those teachers. Thanks also go to Dr. Jocelyn Washburn for her helpful comments on this article.

Notes

1. Low achievers were defined as students who had failed at least one required course in the most recent quarter and who scored below the 33rd percentile on the most recently administered standardized test.
2. To view the chart, go to https://charts.intensive intervention.org/aintervention and look for the strategies listed under "Learning Strategies Curriculum."
3. As the strategies are described here, citations pertaining to the research and theory upon which each strategy was originally founded are included.
4. SWDs = Students with disabilities enrolled in regular English classes.
5. These analyses were calculated after the original research article was published.
6. Based on these and other positive results from other districts, the state started up 40 new Xtreme Reading classrooms by providing professional development workshops to teachers during the summer of 2021 (Leitzel, 2022, personal communication).

7. In fact, the Xtreme Reading Program schedule has been changed for the first year of implementation because repeated experience has shown that the number of strategies scheduled to be taught was too large to fit into the first school year. Two strategies were eliminated for the first year for a teacher: the LINCS Vocabulary Strategy and the Visual Imagery Strategy. They can then be added into the second year of implementation.

References

Afflerbach, P. P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly*, *25*(1), 31–46.

Alvermann, D. E. (2002). Effective literacy instruction for adolescents. *Journal of Literacy Research*, *34*(2), 189–208. https://scholar.google.com/scholar?q=Alvermann+Effective+literacy+instruction+for+adolescents&hl=en&as_sdt=0&as_vis=1&oi=scholart

Alvermann, D. E., & Moore, D. W. (1991). Secondary school reading. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. *2*, pp. 951–983). Longman.

Anderson, B. P. (2009). *Reading instruction at the secondary level, justified or unjustified?: Real strategies for content area reading instruction with applications for social studies*. UNI ScholarWorks. https://scholarworks.uni.edu/grp/272

Baker, E. L. (2004). *Principles for scaling up: Choosing, measuring effects, and promoting the widespread use of educational innovation, CSE Report 634*. CRESST, University of California, LA. https://files.eric.ed.gov/fulltext/ED483401.pdf

Baye, A., Inns, A., Lake, C., & Slavin, R. E. (2018). A synthesis of quantitative research on reading programs for secondary students. *Reading Research Quarterly*, *54*(2), 133–166. https://doi.org/10.1002/rrq.229

Beals, V. L. (1985). The effects of large group instruction on the acquisition of specific learning strategies by learning disabled adolescents. Doctoral dissertation, The University of Kansas. *Dissertation Abstracts International*, *45*(9–A), 2478.

Biancarosa, G., & Snow, C. E. (2004). *Reading next: A vision for action and research in middle and high school literacy*. Carnegie Corporation of New York and Alliance for Excellent Education.

Billingley, B., & Wildman, T. (1988). The effects of prereading activities on the comprehension monitoring of learning disabled adolescents. *Learning Disabilities Research*, *4*(1), 36–44. https://psycnet.apa.orog/record/1989-38191-001

Blachowicz, C., & Fisher, P. (2000). Vocabulary instruction. In M. L. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. *3*, pp. 503–523). Erlbaum.

Boardman, A. G., Roberts, G., Vaughn, S., Wexler, J., Murray, C. S., & Kosanovich, M. (2008). *Effective instruction for adolescent struggling readers: A practice brief*. Center for Instruction. https://files.eric.ed.gov/fulltext/ED521836.pdf

Boudah, D. J. (2018). Evaluation of intensive reading strategies intervention for low-performing adolescents with and without learning disabilities. *Insights into Learning Disabilities*, *15*(2), 145–159. https://files.eric.ed.gov/fulltext/EJ1203396.pdf

Boudah, D. (2022, April). *Applied research on an intensive reading strategies intervention for low-achieving adolescents*. Paper at the Annual Meeting of the American Educational Research Association [virtual].

Bouley, B., Goodson, B., Frye, M., Blocklin, M., & Price, C. (2015). *Summary of research generated by Striving Readers on the effectiveness of interventions for struggling adolescent readers*. Institute of Education Sciences, U.S. Department of Education. Abt Associates, Inc. https://files.eric.ed.gov/fulltext/ED560732.pdf

Brewer, J. (2013). *The word identification strategy: Student materials*. Edge Enterprises, Inc. https://www.edgeenterprisesinc.com

Camera, L. (2019, Oct. 30). Across the board, scores drop in math and reading for U.S. students. *U.S. News & World Report.* https://www.usnews.com/news/education-news/articles/2019-10-30/across-the-board-scores-drop-in-math-and-reading-for-us-students

Carr, E. G., Dunlap, G., Horner, R. H., Koegel, R. L., Turnbull, A. P., Sailor, W., Anderson, J. L., Albin, R. W., Koegel, L. K., & Fox, L. (2002). Positive behavior support: Evolution of an applied science. *Journal of Positive Behavior Interventions*, *4*(1), 4–16. https://doi.org/10/1177/109830070200400102

Chard, D. J., Vaughn, S., & Tyler, B. J. (2002). A synthesis of research on effective interventions for building reading fluency with elementary students with learning disabilities. *Journal of Learning Disabilities*, *35*(5), 386–406. https://doi.org/10.1177/00222194020350050101

Clark, F., Deshler, D. D., Schumaker, J. B., Alley, G. R., & Warner, M. M. (1984). Visual imagery and self-questioning: Strategies to improve comprehension of written material. *Journal of Learning Disabilities*, *17*(3), 145–149. https://doi.org/10.1177/002221948401700304

Corrin, W., Somers, M. A., Kemple, J., Nelson, E., & Sepanik, S. (2009). *The Enhanced Reading Opportunities study: Findings from the second year of implementation* (NCEE 2009–4036). U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences. https://files.eric.ed.gov/fulltext?ED503381.pdf

Day, J. D., Borkowski, J. G., Dietmeyer, D. L., Howsepian, B. A., & Saenz, D. S. (1994). Possible Selves and academic achievement. In L. T. Winegar & J. Valsinen (Eds.), *Children's development within social contexts: Metatheoretical, theoretical and methodological issues* (Vol. *2*). Erlbaum.

Deci, E. L. (2009). Large-scale school reform as viewed from the self-determination theory perspective. *Theory and Research in Education*, *7*(2). https://doi.org/10.1177/1477878509104329

Deshler, D. D., Schumaker, J. B., Bui, Y., & Vernon, D. S. (2006). High schools and adolescents with disabilities: Challenges at every turn. In D. D. Deshler & J. B. Schumaker (Eds.), *Teaching adolescents with disabilities: Accessing the general education curriculum*. Corwin Press.

Deshler, D. D., Schumaker, J. B., Hock, M., & Bulgren, J. A. (2004). *Evaluation of the impact of supplemental literacy interventions in freshman academies*. Technical proposal submitted to the American Institutes for Research.

Deshler, D. D., Schumaker, J. B., & Woodruff, S. K. (2004). Improving literacy skills of at-risk adolescents. In D. S. Strickland & D. E. Alvermann (Eds.), *Bridging the literacy achievement gap grades 4–12* (pp. 174–192). Teachers College Press.

Dillenbourg, P. (2017). *The challenges of scaling-up findings from education research*. Bookings. https://www.brookings.edu/wp-content/uploads/2017/07/meaningful-education-times-uncertainty-essay-12-dillenbourg.pdf

Ellis, E. S. (1992). *The LINCS vocabulary strategy: Instructor's manual*. Edge Enterprises, Inc. https://www.edgeenterprisesinc.com

Ellis, E. S., Deshler, D. D., Lenz, B. K., Schumaker, J. B., & Clark, F. L. (1991). An instructional model for teaching learning strategies. *Focus on Exceptional Children*, *24*(1), 1–14. https://doi.org/10.17161/fec.v23i6.7530 [Reprinted in E. L. Meyen, G. A. Vergason, & R. J. Whelan (Eds.). (1998). *Educating students with mild disabilities* (pp. 151–187). Love Publishing Co.]

Ellis, E. S. & Graves, A.W. (1990). Teaching rural students with learning disabilities: A paraphrasing strategy to increase comprehension of main ideas. *Rural Special Education Quarterly*, *10*, 2–10. https://doi.org/10.1177/875687059001000201

Ellis, E. S., Deshler, D. D., & Schumaker, J. B. (1989). Teaching adolescents with learning disabilities to generate and use task-specific strategies. *Journal of Learning Disabilities*, *22*(2), 108–119, 130.

Faddis, B. J., Beam, M., Maxim, L., Vale Gandhi, E., Hahn, K., & Hale, R. (2011). Portland public schools' striving readers program: Year 5 evaluation report. *RMC Research Corporation*. https://files.eric.ed.gov/fulltext/ED600852.pdf

Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K., Michaels, R., & Shaywitz, S. E. (2015). Achievement gap in reading is present as early as first grade and persists through adolescence. *The Journal of Pediatrics*, *167*(5), 1121–1125. https://doi.org.1.1016/j.jpeds.2015.07.045

Fielding, A., & Schoenback, R. (2003). *An anthology for reading apprenticeship: Building academic literacy*. WestEd. https://WestEd.com

Fisher, J. B., Schumaker, J. B., & Deshler, D. D. (2002). Improving the reading comprehension of at-risk adolescents. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 351–364). Guilford Press.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth-curves analysis. *Journal of Educational Psychology*, *88*(1), 3–17. https://doi.org/10.1037/0022-0663.88.1.3

Fritschmann, N. S., Deshler, D. D., & Schumaker, J. B. (2007a). The effects of instruction in an inference strategy on the reading comprehension skills of adolescents with disabilities. *Learning Disabilities Quarterly*, *30*(4), 244–264. https://doi.org/10.2307/25474637

Fritschmann, N., Schumaker, J. B., & Deshler, D. D. (2007b). *The inference strategy: Instructor's manual*. Edge Enterprises, Inc. https://www.edgeenterprisesinc.com

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239–256 https://doi.org/10.1207/S1532799XSSR0503_3

Fuchs, L., Fuchs, D., & Kazdan, S. (1999). Effects of peer-assisted learning strategies on high school students with serious reading problems. *Remedial and Special Education*, *20*, 309–318. https://doi.org/10.1177/074193259902000507

Gajria, M., & Salvia, J. (1992). The effects of summarization instruction on text comprehension of students. *Exceptional Children*, *58*(6), 508–520. https://journals.sagepub.com/doi/10.1177/001440299205800605

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*(2), 149–164. https://doi.org/10.1177/001440290507100202

Gewertz, C. (2017). National testing landscape continues to shift. *Education Week*, *36*(21), 1. https://www.edweek.org/policy-politics/national-testing-landscape-continues-to-shift/2017/02

Glennan, T. K. Jr., Bodilly, S. J., Galegher, J. R., & Kerr, K. A. (2004a). *Expanding the reach of education reforms: Perspectives from leaders in the scale-up of educational interventions*. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/monographs/2004/RAND_MG248.pdf

Glennan, T. K. Jr., Bodily, S. J., Galegher, J. R., & Kerr, K. A. (2004b). *Expanding the reach of education reforms: What have we learned about scaling up educational interventions?* RAND Corporation. https://www.rand.org/pubs/research_briefs/RB9078.html

Hammill, D. D., Wiederholt, J. L., & Allen, E. A. (2006). TOSCRF-2: Test of silent contextual reading fluency. Examiner's manual. Pro-Ed.

Hempenstall, K. (2021). *Teaching reading in secondary schools: Some theoretical and practical issues*. National Institute for Direct Instruction. https://www.dropbox.com/sh/olxpifutwcgvg8j/AAAIKplGUIX1-XqTgEnXBC4L1a/PDFs?dl=0&preview=Teaching+reading+in+secondary+schools.pdf

Henderson, A. J., & Shores, R. E. (1982). How learning disabled students' failure to attend to suffixes affects their oral reading performance. *Journal of Learning Disabilities*, *15*, 178–182. https://doi.org/10.1177/002221948201500311

Harris, M., Schumaker, J. B., & Deshler, D. D. (2008). *The word mapping strategy: Instructor's manual*. Edge Enterprises, Inc. https://www.edgeenterprisesinc.com

Harris, M. L., Schumaker, J. B., & Deshler, D. D. (2011). The effects of strategic morphological analysis instruction on the vocabulary performance of secondary students with and without disabilities. *Learning Disability Quarterly*, *34*(1), 1–15. https://doi.org/10.1177/073194871103400102

Hock, M. F., Deshler, D. D., & Schumaker, J. B. (1993). Learning strategy instruction for at-risk and learning-disabled adults: The development of strategic learners through apprenticeship. *Preventing School Failure*, *38*(1), 43–49. https://foi.otf/10.1080/1045988X.1993.9944294

Hock, M. F., Deshler, D. D., & Schumaker, J. B. (2005). Enhancing student motivation through the pursuit of possible selves. In C. Dunkel & J. Kerpelman (Eds.), *Possible selves: Theory, research, and application* (pp. 205–221). Nova Science Publishers.

Hock, M. F., Pulvers, K. A., Deshler, D. D., & Schumaker, J. B. (2001a). The effects of an after-school tutoring program on the academic performance of at-risk and students with learning disabilities. *Remedial and Special Education*, *22*(3), 172–186. https://doi.org/10.1177/074193250102200305

Hock, M. F., Schumaker, J. B., & Deshler, D. D. (2001b). The case for strategic tutoring. *Educational Leadership*, *58*(7), 50–52.

Hock, M. F., Schumaker, J. B., & Deshler, D. D. (2003). *Possible selves: Nurturing student motivation*. Edge Enterprises, Inc. http://www.edgeenterprisesinc.com

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, *2*(2), 127–160. https://doi.org/10.1007/BF00401799

Hougen, M. (2015). *Evidence-based reading instruction for adolescents in Grades 6–12* (CEEDAR Document No IC-13). The CEEDAR Center. https://ceedar.education.ufl.edu/wp-content/uploads/2015/05/IC-13_FINAL_05-26-15.pdf

Hughes, C. A., Deshler, D. D., Ruhl, K. L., & Schumaker, J. B. (1993). Test-taking strategy instruction for adolescents with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders*, *1*(3), 189–198. https://doi.org/10.1177/106342669300100307

Hughes, C. A., & Schumaker, J. B. (1991). Test-taking strategy instruction for adolescents with learning disabilities. *Exceptionality*, *2*, 205–221. https://doi.org/10.1080/09362839109524784

i3 Community. (2017). *Scaling up evidence-based practices: Strategies from investing in innovation*. Westat. https://oese.ed.gov/files/2022/03/Scaling_Up_Evidence_Based_Practices_i3.pdf

Islam, S., & Santoso, E. (2019). The effectiveness of using authentic texts in teaching reading comprehension. *ETERNAL*, *4*(2), 181–199. https://doi.org/10.24252/Eternal.V42.2018.A3

Jitendra, A. K., Edwards, L. L., Sacks, G., & Jacobson, L. A. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children*, *70*(3), 299–322. https://doi.org/10.1177/001440290407000303

Johnson, D. W., & Johnson, R. T. (1990). Social skills for successful group work. *Educational Leadership*, *47*(4), 29–33. https://eric.ed.gov/?id=EJ400495

Johnson, D. W., Johnson, R. T., & Holubed, E. J. (1994). *The nuts and bolts of cooperative learning*. Interaction Book Company.

Kamil, M. L. (2003). *Adolescents and literacy: Reading for the 21st century*. Alliance for Excellent Education. https://media.carnegie.org/filer_public/ff/cc/ffcc9965-bed2-406e-8e6e-907e09b8cb8c/ccny_grantee_2003_adolescents.pdf

Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). Improving adolescent literacy: Effective classroom and intervention practices: A practice guide *(NCEE #2008-4027)*. U.S. Department of Education, Institute of Education Sciences. http://ies.ed.gov/ncee/wwc

Kemple, J. 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). U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences. http://ies.ed.gov/ncee/wwc

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*(4), 294–303. https://doi.org/10.1037/0003-066X.49.4.294

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kline, F. M., Schumaker, J. B., & Deshler, D. D. (1991). Development and validation of feedback routines for instructing students with learning disabilities. *Learning Disability Quarterly*, *14*(3), 191–207. https://doi.org/10.2307/1510849

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, *95*(1), 3. http://www.ciera.org/library/reports/inquiry-2/2-008/2-008.html.

Lang, L., Torgesen, J., Vogel, W., Chanter, C., Lefsky, E., & Petscher, Y. (2009). Exploring the relative effectiveness of reading interventions for high school students. *Journal of Research on Educational Effectiveness*, *2*(2), 149–175. https://doi.org/10.1080/19345740802641535

Lenz, B. K., & Hughes, C. A. (1990). A word identification strategy for adolescents with learning disabilities. *Journal of Learning Disabilities*, *23*(3), 149–158. https://doi.org/10.1177/002221949002300304

Lenz, B. K., Schumaker, J. B., Deshler, D. D., & Beals, V. L. (1984). *The word identification strategy: Instructor's manual*. University of Kansas Center for Research on Learning. https://sim.ku.edu/word-identification-strategy

Lenz, B. K., Schumaker, J. B., Deshler, D. D., & Boudah, D. (1994). *The unit organizer routine: Instructor's manual*. Edge Enterprises, Inc. https://www.edgeenterprisesinc.com

Lenz, B. K., Schumaker, J. B., Deshler, D. D., & Bulgren, J. A. (1998). *The course organizer routine: Instructor's manual*. Edge Enterprises, Inc. https://www.edgeenterprisesinc.com

Lexile. (2022). *Lexile framework for reading*. https://www.lexile.com/educations/measuring-growth-with-lexile/college-and-career-readiness/

Ludwig, C. I. (2019). *Why aren't secondary schools making reading instruction a priority?* RTS Success. https://rtssuccess.com/why-arent-secondary-schools-making-secondary-reading-instruction-a-priority/

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading tests* (4th edn.). The Riverside Publishing Company.

Malone, L. D., & Mastropieri, M. A. (1992). Reading comprehension instruction: Summarization and self-monitoring training for students with learning disabilities. *Exceptional Children*, *58*(3), 270–279. https://doi.org/10.1177/001440299105800309

Markus, H., & Nurris, P. (1986). Possible selves. *American Psychologist*, *41*, 954–969. https://psycnet.apa.org/doiLanding?doi=10.1037%2F0003-066X.41.9.954

Nation, I.S.P. (1990). *Teaching and learning vocabulary*. Newbury House.

National Center for Educational Statistics. (2015). *NAEP report card: 2019 NAEP reading assessment*. U.S. Department of Education, Institute of Education Sciences. https://www.nationsreportcard.gov/highlights/reading/2015/

National Center for Educational Statistics. (2019). *NAEP report card: 2019 NAEP reading assessment*. U.S. Department of Education, Institute of Education Sciences. https://www.nationsreportcard.gov/highlights/reading/2019/

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010a). *Common core state standards*.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010b). *Common core state standards for English language arts & literacy in history/social studies, and technical subjects*. http://www.corestandards.org/wp-content/uploads/ELA_Standards1.pdf

National Reading Panel. (2000). *Teaching children to read*. http://www.nationalreadingpanel.org

O'Shea, L. J., Sindelar, P. T., & O'Shea, D. J. (1985). The effects of repeated readings and attentional cues on reading fluency and comprehension. *Journal of Reading Behavior*, *17*, 129–142. https://doi.org/10.1080/10862968509547535

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, *40*, 357–367. http://doi.org/10.1016/j.beth.2008/10.006

Pressley, M., Borkowski, J. G., and Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate meta-cognition and knowledge. In R. Vasta & G. Whitehurst (Eds.), *Annals of child development* (Vol. *4*, pp. 89–129). Greenwich, CT: JAI Press.

Pressley, M., Borkowski, J. G., & Schneider, W. (1987a). Cognitive strategies: Good strategy users coordinate meta-cognition and knowledge. In R. Vasta & G. Whitehurst (Eds.), *Annals of child development* (Vol. *4*, pp. 89–129). JAI Press.

Pressley, M., Johnson, C. J., & Symons, S. (1987b). Elaborating to learn and learning to elaborate. *Journal of Learning Disabilities*, *20*, 76–91. https://doi.org/10.1177/002221948702000202

Pressley, M., & McCormick, C. (1995). *Cognition, teaching, and assessment*. Harper Collins.

Putnam, L. (1988). *An investigation of the curricular demands in secondary mainstream classrooms containing students with mild handicaps*. Doctoral dissertation, University of Kansas. Dissertation Abstracts.

Putnam, L. M., Deshler, D. D., & Schumaker, J. B. (1992). The investigation of setting demands: A missing link in learning strategy instruction. In L. Meltzer (Ed.), *Strategy assessment and instruction for students with learning disabilities: From theory to practice* (pp. 325–351). Pro-Ed.

Quigley, A., & Coleman, R. (2020). *Guidance report: Improving literacy in secondary schools*. Evidence for Learning.

Quint, J., Bloom, H. S., Black, A. R., Stephens, L., & Akey, T. M. (2005). *The challenge of scaling up educational reform: Findings and lessons from first things first (final report)*. MDRC. https://eric.ed.gov/?od=ED485680

Reynolds, D. (2020). Of research reviews and practice guides: Translating rapidly growing research on adolescent literacy into updated practice recommendations. *Reading Research Quarterly*, *56*(3), 401–414. https://doi.org/10.1002/rrq.314

Reynolds, D. (2021). Updating practice recommendations: Taking stock of 12 years of adolescent literacy research. *Journal of Adolescent & Adult Literacy*, *65*(1), 37–46. https://doi.org/10.1002/jaa1176

Roehler, L. R., & Duffy, G. G. (1984). Direct explanation of comprehension processes. In G. G. Duffy, L. R. Duffy, & J. Mason (Eds.), *Comprehension instruction: Perspectives and suggestions* (pp. 265–280). Longman.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press.

Safer, N., & Fleischman, S. (2005). How student progress monitoring improves instruction. *Educational Leadership*, *62*(5), 81–83. https://www.quop.de/fileadmin/literatur/NancySafer_ResearchMatters.pdf

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook for research in early literacy* (pp. 97–110). Guilford Press.

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

Schmidt, J. L., Deshler, D. D., Schumaker, J. B., & Alley, G. R. (1988/89). Effects of generalization instruction on the written language performance of adolescents with learning disabilities in the mainstream classroom. *Reading, Writing, and Learning Disabilities*, *4*(4), 291–309.

Schumaker, J. B. (2008). *The hidden results of a scaled-up reading strategy intervention. Research Report #66*. University of Kansas Center for Research on Learning.

Schumaker, J. B., & Deshler, D. D. (1984). Setting demand variables: A major factor in program planning for the LD adolescent. *Topics in Language Disorders*, *4*(2), 22–40. https://doi.org/10.1097/00011363-198403000-00006

Schumaker, J. B., & Deshler, D. D. (1992). Validation of learning strategy interventions for students with learning disabilities: Results of a programmatic research effort. In B.Y.L. Wong (Ed.), *Contemporary intervention research in learning disabilities: An International perspective* (pp. 22–46). Springer-Verlag.

Schumaker, J. B., & Deshler, D. D. (2006). Teaching adolescents to be strategic learners. In D. D. Deshler & J. B. Schumaker (Eds.), *Teaching adolescents with disabilities: Accessing the general education curriculum* (pp. 121–156). Corwin Press.

Schumaker, J. B., Deshler, D. D., Hock, M., & Bulgren, J. A. (2015). *The Xtreme reading program*. University of Kansas Center for Research on Learning. https://sim.ku.edu/xtreme-reading

Schumaker, J. B., Deshler, D. D., Nolan, S. M., & Alley, G. R. (1994). *The self-questioning strategy: Instructor's manual*. University of Kansas Center for Research on Learning. https://sim.ku.edu/self-questioning-strategy

Schumaker, J. B., Deshler, D. D., Woodruff, S., Hock, M. F., Bulgren, J. A., & Lenz, B. K. (2006). Reading strategy interventions: Can literacy outcomes be enhanced for at-risk adolescents? *Teaching Exceptional Children*, *38*(3), 64–69. https://doi.org/10.1177/004005990603800310

Schumaker, J. B., & Deshler, D. D. (2007). *The inference strategy: Student materials*. Edge Enterprises, Inc.

Schumaker, J. B., Deshler, D. D., Zemitzch, A., & Warner, M. M. (1993). *The visual imagery strategy: Instructor's manual*. University of Kansas Center for Research on Learning. https://sim.ku.edu/visual-imagery-strategy

Schumaker, J. B., Knight, J., & Deshler, D. D. (2007). *Fundamentals of paraphrasing and summarizing: Instructor's manual*. Edge Enterprises, Inc. https://www.edgeenterprisesinc.com

Schumaker, J. B., Denton, P. H., & Deshler, D. D. (1984). *The paraphrasing strategy: Instructor's manual*. University of Kansas Center for Research on Learning Disabilities. https://sim.ku.edu/paraphrasing-strategy

Shanahan, T. (2018, May 19). Comprehension skills or strategies: Is there a difference and does it matter? *Shanahan on Literacy*. https://www.shanahanonliteracy.com/blog/comprehension-skills-or-strategies-is-there-a-difference-and-does-it-matter?fbclid=IwAR3eUXpKq9CfDa3vp1jUbiihJT7IjU8SKNZCKlaSoT97mHD7Po3UI_TmrSY#sthash.j0VgWI3M.NTFu4ChG.dpbs

Shannon, L., & Grant, B. J. (2015). *A final report for the evaluation of the Achieve3000 programs*. Magnolia Consulting.

Shrank, F. A., & Wendling, B. J. (2018). *The Woodcock-Johnson IV: Tests of cognitive abilities, tests of oral language, tests of achievement*. Houghton Mifflin Harcourt - Riverside.

Siegel, S., & Castellan Jr., N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd edn.). McGraw Hill.

Simmonds, E.P.M. (1992). The effects of teacher training and implementation of two methods of improving the comprehension skills of students with learning disabilities. *Learning Disabilities Research and Practice*, *7*(4), 194–198. https://psychnet.apa.org/record/1993-15522-001

Sirles, C. (1997). *Root awakenings: Vocabulary development using classical word roots*. Stripes Publishing.

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*(3), 290–322. https://doi.org/10.1598/RRQ.43.3.4

Slosson, R. L. (1990). *Slosson oral reading rest-revised*. Slosson Educational.

Snow, C. (2002). *Reading for understanding: Toward an R & D program in reading comprehension*. RAND.

Somers, M. A., Corrin, W., Sepanik, S., Salinger, T., Levin, J., & Zmach, C. (2010). The Enhanced Reading Opportunities study final report: The impact of supplemental literacy courses for struggling ninth-grade readers *(NCEE 2010–4021)*. U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences. https://eric.ed.gov/?id=ED511811

Spear-Swerling, L., & Sternberg, R. (2002). What science offers teachers of reading. *Learning Disabilities Research & Practice*, *16*(1), 51–57. https://www.readingrockets.org/article/what-science-offers-teachers-reading

Spencer, A. (2000). *Morphological theory: An introduction to word structure in generative grammar*. Blackwell Publishers.

Sprague, K., Zaller, C., Kite, A., & Hussar, K. (2012). *Springfield-Chicopee School Districts Striving Readers (SR) Program: Final report years 1–5: Evaluation of implementation and impact*. Prepared for Office of Elementary and Secondary Education, U.S. Department of Education Institute of Education Appendix 13 Sciences, U.S. Department of Education. https://eric.ed.gov/?id=ED600926

Sprick, R. S., & Garrison, M. (1998). *CHAMPS: A proactive and positive approach to classroom management*. SoprisWest.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://doi.org/10.1598/RRQ.21.4.1

Stenner, A. J., Sanford-Moore, E., & Williamson, G. L. (2012, Oct.). *The Lexile Framework for reading quantifies the reading ability needed for "college & career readiness."* Metametrics Research Brief.

Stevens, R. J., & Durkin, S. (1992). *Using student team reading and student team writing in middle schools: Two evaluations* (Report No. 36). Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students.

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research for students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, *68*(3), 276–321. https://doi.org/10.3102/00346543068003277

Swanson, H. L., Hoskyn, M., & Lee, C. (1999). *Interventions for students with learning disabilities: A meta-analysis of treatment outcomes*. Guilford Press.

Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, C. M., Scott, T., Liaupsin, C., Sailor, W., Turnbull, A. P., Turnbull, H. R., III, Wickham, D., Reuf, M., & Wilcox, B. (2000). Applying positive behavioral support and functional behavioral assessment in schools. *Journal of Positive Behavioral Interventions*, *2*, 131–143. https://doi.org/10.1177/109830070000200302

Torgesen, J. K., Rashotte, C. A., & Alexander, A.W. (2001). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 333–355). York Press.

Tralli, R., Colombo, B., Deshler, D. D., & Schumaker, J. B. (1996). The strategies intervention model: A model for supported inclusion at the secondary level. *Remedial and Special Education*, *17*(4), 204–216. https://doi.org/10.1177/074193259601700403 [Reprinted in G. Bunch & A. Valeo (Eds.). (1997). Inclusion: Recent research. Inclusion Press; and in D. D. Deshler, J. B. Schumaker, K. R. Harris, & S. Graham (Eds.). (1999). *Teaching every adolescent every day: Learning in diverse schools and classrooms*. Brookline Books.]

Vaughn, S. (2019). *Providing reading interventions for students in grades 4–9: Educator's practice guide*. Institute of Education Sciences. https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/WWC-practice-guide-reading-intervention-full-text.pdf#page=11

Vaughn, S., & Fletcher, J. M. (2012). Response to intervention with secondary school students with reading difficulties. *Journal of Learning Disabilities*, *45*(3), 244–256. https://doi.org/10.1177/0022219412442157

Vaughn, S., Roberts, G., Klingner, J. K., Swanson, E. A., Boardman, A., Stillman-Spisak, S. J., & Leroux, A. J. (2013). Collaborative strategic reading: Findings from experienced implementers. *Journal of Research on Educational Effectiveness*, *6*(2), 137–163. https://doi.org/10.1080/19345747.2012.741661

Vernon, D. S., Schumaker, J. B., & Deshler, D. D. (1993). *The SCORE Skills: Social skills for use in cooperative teaming structures*. Edge Enterprises, Inc. http://www.edgeenterprisesinc.com

Vernon, D. S., Deshler, D. D., & Schumaker, J. B. (2000). *Talking together: Instructor's manual*. Edge Enterprises, Inc. http://www.edgeenterprisesinc.com

Warner, M. M., Schumaker, J. B., Alley, G. R., & Deshler, D. D. (1980). Learning disabled adolescents in the public schools: Are they different from other low achievers? *Exceptional Education Quarterly*, *1*(2), 27–35. https://doi.org/10.1177/074193258000100207 [Reprinted in American Library Association. (1982). *Mainstreamed library: Issues, ideas, innovations*.]

Williams, K. T. (2001). *GRADE: Group Reading Assessment & Diagnostic Evaluation*. American Guidance Service.

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its interventions. *Scientific Studies of Reading*, *5*(3), 211–239. https://doi.org/10.1207/S1532799XSSR0503_2

Woodruff, S., Schumaker, J. B., & Deshler, D. D. (2002). *The effects of an intensive reading intervention on the decoding skills of high school students with reading deficits*. (Research Report No. 15). University of Kansas Center for Research on Learning.

About the Author

Jean Schumaker, PhD, SEP, is Professor Emeritus in the Departments of Special Education and Applied Behavioral Science at the University of Kansas. She is also President of Edge Enterprises, Inc., a research and publishing company focusing in the area of special education. She is a certified trauma therapist. She received her PhD from the University of Kansas in Developmental and Child Psychology in 1976. She has spent the last 50 years studying the problems of adolescents and developing educational interventions for them.