

The Effects of Theoretically Different Instruction and Student Characteristics on the Skills of Struggling Readers

Patricia G. Mathes, Carolyn A. Denton, Jack M. Fletcher, Jason L. Anthony, David J. Francis, and Christopher Schatschneider

Perhaps the most important responsibility of educators in the primary grades is to ensure that all students become competent readers. The degree of success in becoming a competent reader typically is established in the early grades (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Juel, 1988; Torgesen & Burgess, 1998). Unless effective instructional practices are used in this critical period, the inequities that commonly divide our students are likely to continue (Snow, Burns, & Griffin, 1998).

During the previous 25 years there have been numerous studies focusing on the prevention of reading problems with young students. Converging evidence from these studies suggests that early instruction can be effective in preventing reading problems for many students (see Denton & Mathes, 2003; Snow et al., 1998) and that real schools and teachers can implement this instruction (e.g. Clay, 1993; Mathes & Torgesen, 1998). Further, it appears that without effective early instruction, which may require supplemental instruction, initial reading difficulties may eventually be compounded as students fall further and further behind their peers and habituate ineffective strategies for coping with reading failure (Clay, 1987; Stanovich, 1986). In other words, students may “learn to be learning disabled” (Clay, 1987, p. 155).

Pedagogical Framework

Beyond understanding the role that early intervention can play in reducing reading difficulties, we also have a greater understanding of the critical content for which struggling readers must gain ownership if they are to become competent readers. This content provides a framework to guide pedagogical decisions and includes provisions for teaching phonemic awareness, phonemic decoding skills, fluency in word recognition and text processing, construction of meaning, vocabulary, spelling, and writing (see Foorman & Torgesen, 2001; National Institute of Child Health and Human Development [NICHD], 2000; Pressley, 1998; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001; Snow et al., 1998). Likewise, it appears that instruction in these areas needs to be explicit. By explicit we mean that students are not required to infer new knowledge, but rather that new knowledge is shared directly. It also appears that for some students, instruction must be intensive in order to facilitate adequate reading development. By intensive instruction we mean that students are highly engaged in learning critical content and that the ratio of teachers to students is relatively small.

Unanswered Questions

What is less clear is exactly how much instruction must occur, how contextualized skills instruction needs to be, and the level of intensity at which it must occur in order for struggling readers to succeed. It is logical to assume that high-quality classroom instruction is a primary factor in determining success. The findings of several studies confirm that appropriate classroom-level instruction in the primary grades can dramatically reduce the prevalence of reading problems (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Mathes, Howard, Allen, & Fuchs, 1998; Mathes, Torgesen, & Allor, 2001). However, even when classroom instruction is of high quality, approximately 5% to 7% of students do not meet benchmarks associated with reading proficiency in the early grades (Denton & Mathes, 2003; Mathes & Denton, 2002; Torgesen, 2000). For these students, it would appear that more intensive instruction is required.

There are many examples of successful individual and small-group intense interventions. For example, many struggling first-grade readers who receive one-on-one tutoring in the Reading Recovery program (Clay, 1993) routinely make substantial progress, enabling them to accurately read connected text at a level of difficulty similar to that of their more able peers (i.e., Gomez-Bellenge, Rogers, & Fullerton, 2003). Other interventions using a variety of methods show that individual and small-group tutorial programs effectively reduce the number of struggling readers (O'Connor, 2000; Simmons, Kame'enui, Stoolmiller, Coyne, & Harn, 2003; Torgesen et al., 1999; Torgesen, Rashotte, Alexander, Alexander, & MacPhee, 2003; Vellutino et al., 1996). While these interventions vary in how instruction is provided and in the amount of emphases placed on various content, each one generally reduces the number of struggling readers to 4.5% or less of the school population (Mathes & Denton, 2002; Torgesen et al., 2003).

What has not yet been investigated is the effect of providing high-quality classroom-level instruction in tandem with intense supplemental

small-group interventions. Thus, while it is known that either enhanced classroom-level instruction or individual and small-group interventions significantly reduce the number of struggling readers, the effect of both practices together is not known. Further, while it is clear that many students at risk for reading failure need intense early reading instruction composed of content that research suggests is critical, it has yet to be determined whether individual characteristics of struggling readers can be identified that will assist practitioners in matching interventions to learner needs. To date, few researchers have entertained the idea that there may be an interaction between learner characteristics and the efficacy of specific approaches.

What is known is that individual students vary in their development of reading-related abilities. Thus, it is reasonable to speculate that students vary in which aspects of reading instruction are most critical. Generally speaking, some of the best validated predictors of later reading ability include letter knowledge (Simmons et al., 2003; Vellutino, Scanlon, & Jaccard, 2003), phonological awareness (Foorman et al., 1998; Torgesen et al., 1999; Vellutino et al., 2003), rapid naming of letters (Simmons et al., 2003; Vellutino et al., 2003), verbal memory (Vellutino et al., 2003), and oral language, including vocabulary (Storch & Whitehurst, 2002). To date, no researcher has changed the type of instruction provided to children with varying characteristics to determine if response to instruction is affected by children's individual characteristics.

In conducting the current research, we investigated the effectiveness of combining enhanced classroom instruction and intense supplemental intervention for struggling readers in first grade. Further, we explored the efficacy of two supplemental interventions derived from diverse theoretical foundations that we believe are widely embraced, at least implicitly, in today's schools, examining them in terms of interactions with child characteristics and academic outcomes.

One intervention, *Proactive Reading*, is derived from the model of Direct Instruction (Carnine, Silbert, Kame'enui, & Tarver, 2004; Engelmann, 1997; Engelmann & Carnine, 1982; Kame'enui & Simmons, 1990), which has its foundation in behavioral theory (Becker, 1973; Skinner, 1953) but goes beyond Skinnerian behaviorism to include teacher communications, student responses, and knowledge forms as elements for consideration when designing instruction (Engelmann & Carnine). The central characteristic of reading instruction based on behavioral theory is a focus on directly observable reading behaviors rather than on making inferences about the learner's cognitive processing (Pressley & McCormick, 1995). Behaviorism is a positive approach to learning in that the teacher supports the development of behaviors (such as the accurate and fluent application of reading skills) through positive reinforcement (such as praising and rewarding students for doing what is being taught). Skills instruction based on a behaviorist perspective is designed to be systematic; simpler skills are mastered before progressing to more complex skills. The role of the teacher in a behaviorist model is to teach content and model skills and then to provide practice and reinforcement for mastering those skills.

Proactive Reading includes these features but goes beyond Skinnerian behaviorism to include teacher communications, student responses, and knowledge forms as elements for consideration when designing instruction. In *Proactive Reading*, the tasks associated with fluent, meaningful reading were analyzed and systematically arranged into a scope and sequence intended to reduce student confusion and support successful learning. This was also intended to ensure that each child was gradually gaining cumulative knowledge and skills, resulting in ever-greater ability to read increasingly more complex text. From this scope and sequence daily lessons were derived. The result was that students learned phonetic elements in isolation before applying them strategically to words and practiced decoding words in isolation before reading decodable connected text and

applying comprehension strategies. Over time, the types of words read increased in complexity; stories became increasingly longer and story lines more complex. The teacher's job was to deliver this instruction as prescribed, following a daily lesson plan in which specific wording and expected student responses were provided. This wording was designed to communicate to children only critical attributes of new content without introducing possibly confusing information or overloading the student with too much information. Following a behavioral approach, lessons provided for frequent reinforcement on both an interval and intermittent schedule. This reinforcement included verbal praise provided as children performed tasks well, check marks on a mastery record, and stickers at the end of tasks and lessons. Likewise, when errors occurred, teachers retaught basic facts or scaffolded more complex content and then provided two to three additional practice opportunities for that fact or content.

The other supplemental intervention, *Responsive Reading*, is aligned with cognitive theory, following the model of cognitive strategy instruction (see Harris & Pressley, 1991), a form of cognitive apprenticeship (Brown, Collins, & Duguid, 1989; Rogoff, 1990; Rojewski & Schell, 1994). This model characterizes learning in terms of the acquisition of problem-solving strategies through a process of modeling, guided practice, coaching, scaffolding, and fading (Brown et al.; Rojewski & Schell). It is the role of the teacher to make his or her own knowledge explicit and to model strategies and then to coach and scaffold the learners as they apply these concepts and strategies in authentic activity. Ultimately, students are empowered to apply strategies independently. This model is also related to the theories of Vygotsky (1978), in that the social nature of learning is critical, and the teacher creates situations for learning that are sufficiently challenging to engage the learner in problem solving but not so challenging that the learner cannot achieve success (Rogoff). In their description of cognitive strategy instruction, Harris and Pressley discussed the development of preskills

necessary for application of problem-solving strategies. Responsive Reading follows a pattern of explicit instruction in essential preskills and the modeling of strategies, which is then followed by application of these skills and strategies in reading and writing authentic text with teacher support and scaffolding. Responsive Reading is similar to the guided reading model (see Fountas & Pinnell, 1996) in its emphasis on modeling, prompting, and scaffolding within the learner's zone of proximal development (Vygotsky, 1978). It differs from that approach in terms of the explicit nature of instruction in essential preskills and the teaching of word identification to reflect current research findings in beginning reading (e.g., Snow et al., 1998).

The objectives of daily instruction within Responsive Reading were determined by the observed needs of students. Thus, there was no predetermined scope and sequence. The Responsive Reading teacher attended to behaviors that indicated the constructions each student was building up and responded contingently, providing scaffolds to help the student perform more effectively and promoting increasing independence over time. The Responsive Reading teacher had a dual role, providing explicit instruction in reading skills such as phonemic awareness and phonics but also serving as a coach to the student, directing the student's development of useful strategies while the student was engaged in reading and writing connected text. As students read stories chosen to reflect their ability and wrote words and sentences, their teacher observed what was easy for them and what types of errors were made, and he or she planned instruction accordingly. Thus, while both interventions were based on the same pedagogical content, the way instruction was delivered in the two approaches was distinctly different, reflecting the different theoretical orientations from which they were derived.

While different in their theoretical underpinnings, both interventions were comprehensive, integrated approaches to reading instruction that incorporated content deemed

critical in recent reports to promote successful reading acquisition (NICHD, 2000; Pressley, 1998; Snow et al., 1998). Thus, the two approaches provided instruction in phonemic awareness, alphabetic knowledge and skills, and application of this knowledge to words and text, and they engaged students in making meaning from what they had read. However, they differed in the way this content was taught.

Research Hypotheses

In the current research, we asked two overarching questions. First, we asked if small-group supplemental intervention derived from either behavioral or cognitive theory provided in addition to research-based reading instruction was more effective than such instruction alone in promoting greater academic growth among struggling first-grade readers. Second, we asked if certain child characteristics known to be critically important for reading acquisition could differentially predict children's responses to the interventions, including phonological awareness, rapid naming of letters, and vocabulary. Specifically, we tested four hypotheses:

1. We hypothesized that small-group reading instruction, in the form of the Proactive or Responsive Reading intervention, provided in addition to the classroom reading program, would be, on average, more effective than high-quality classroom reading instruction alone for students at risk for reading failure. This prediction stems from the finding that interventions at different places on a direct instruction/cognitive apprenticeship continuum accelerated reading development. Therefore, adding intervention to classroom instruction should be more effective than classroom instruction alone.
2. Likewise, we predicted that despite differences in content emphasis and level of contextualization, the two supplemental interventions would be, on average, comparably effective because both incorpo-

rated instruction in critical components of learning to read.

3. We also hypothesized that because the interventions would accelerate development, the reading performance of struggling first-grade readers who received either supplemental intervention would approach the level of performance of their normally developing peers. We based this hypothesis on the fact that studies providing either enhanced core reading or individual or small-group reading intervention resulted in significantly greater numbers of children attaining normal reading levels (e.g., Denton & Mathes, 2003; Mathes & Denton, 2002). Thus, we speculated that coupling core and small-group reading instruction would result in even more children attaining normal reading levels.
4. Finally, based on the literature on predictors of response to intervention and reading outcomes, we hypothesized that initial phonological awareness, rapid automatic letter naming, and vocabulary would differentially predict individuals' responses to the two supplemental interventions. Based on differences in expected time devoted to different components of reading in the two interventions, we expected the Proactive approach to be more effective for children struggling with phonological processing and the Responsive approach to be more effective for children struggling with vocabulary development.

Method

Participants

Schools. This research was conducted in six U.S. schools in a large urban school district in Texas. We selected these schools because they had been designated as relatively high-performing schools in reading by the state's department of education and the school district, which we used as an indicator of a successful

classroom reading program. One school was designated as *acceptable* (with high reading scores), two were identified as *recognized*, and three were designated as *exemplary* when the study began. Performance on a nationally normed, group-administered reading achievement test at the end of first grade indicated that the average reading performance in each of the six schools was above the national average. As we did not have the resources to provide extensive classroom-level intervention, selection of these schools helped ensure that we were examining the effect of quality classroom reading instruction with and without supplemental intervention. None of these schools were eligible for Title I (a federally funded program for at-risk students), and all served diverse student populations in terms of ethnicity and socioeconomic status.

Students. During each of two years, we identified within these schools a sample of first graders who showed significant risk for reading difficulties. In order to determine which students were at risk for reading difficulty, classroom teachers and our research team screened all students at the end of kindergarten within the six participating schools using the kindergarten screening portion of the Texas Primary Reading Inventory (TPRI). At the beginning of the first-grade year we screened any students entering the school for the first time with the first-grade TPRI screen.

The kindergarten screen of the TPRI was derived from a large longitudinal study of students in kindergarten through grade 2 (Schatschneider, Francis, Foorman, Fletcher, & Mehta, 1999). At kindergarten, it consists of a two-minute assessment of letter-sound knowledge and phonological awareness (blending sounds). However, this screen is designed to maximize the probability that students with risk characteristics would not be missed (i.e., false negative errors), resulting in overidentification of risk status (false positive errors). In order to accurately discriminate at-risk status, we followed the TPRI screening with the administration of the Woodcock-Johnson III

(W–J III) Word Identification subtest and the text reading subtest of the Observation Survey of Early Literacy Achievement, eliminating any student who could read five words or more on the Woodcock–Johnson or who could read texts designated as Level D or higher (Fountas & Pinnell, 1999) with at least 90% accuracy. Further, we collected a one-minute oral reading sample on an end-of-first-grade passage and only included children reading five or fewer words correctly per minute.

All students served in regular education classes were eligible for the study, including students who qualified for special education based on the identification of a learning disability, speech or language impairment, or “other health impairment.” We excluded students with limited English proficiency who were served in bilingual classrooms and students served primarily in self-contained special education classes.

Once identified, all students designated as at risk *within* a school were randomly assigned to one of three conditions: enhanced classroom + Proactive Reading, enhanced classroom + Responsive Reading, or enhanced classroom only. In addition, we identified a sample of typically achieving readers by randomly selecting them from among all students in the same classrooms who evidenced no risk for reading problems. The purpose of this typically achieving group was to provide a benchmark of typical reading development in these classrooms.

To increase sample size, the study was conducted over two successive school years with two cohorts of students. In total, our sample included 92 students in the Proactive intervention group, 92 students in the Responsive intervention group, 114 students in the at-risk enhanced classroom condition, and 101 students who were typically achieving. After the effects of attrition, 78 Proactive Reading students, 83 Responsive Reading students, 91 at-risk students who received quality classroom instruction only, and 94 typically achieving students were assessed at posttest.

Table 1 summarizes the demographic information and educational status information

for all participants because statistical analyses included all children. Given that the small number of Asian American children precluded treating them as a separate group in analyses and given that these children’s language and literacy scores closely paralleled those of the Caucasian children, the Asian American and Caucasian groups were combined for analyses that sought to evaluate ethnicity effects prior to the main hypothesis testing. Most notably, no statistically significant differences among the at-risk groups were detected for any of the demographic or educational status variables.

Intervention Teachers. We employed six certified teachers to provide the intensive supplemental instruction. Three of the teachers taught Proactive Reading and three taught Responsive Reading. Each teacher taught at two different schools during each school day, enabling us to place both a Proactive and a Responsive teacher in each school so that school effects and intervention teacher effects were not confounded. Four of the six teachers held master’s degrees, and several had teaching certifications in multiple areas. Five were certified in elementary general education, two in special education, three in English as a second language, one in early childhood education, and two as reading specialists. Two of the teachers also held certificates in educational administration. The mean years of teaching experience for the six teachers prior to the onset of the study was nine years, with a range from 3 to 22 years. All six teachers were experienced at teaching primary-grade students. The same six teachers delivered the Proactive and Responsive interventions during the two years of the study.

Classroom Teachers. Thirty first-grade classroom reading teachers from the six schools participated in this research across two years. Sixteen of these teachers participated in both years of the study. All teachers used one of two basal reading series adopted by the district and selected by their respective schools. Both of these basal programs provided guidance for delivering a comprehensive reading curriculum

Table 1
Student Demographic Information by Group

	Proactive			Responsive			Enhanced Classroom			Typically Achieving		
	n	%	M (SD)	n	%	M (SD)	n	%	M (SD)	n	%	M (SD)
Age in months	92	—	78 (4.9)	92	—	78 (4.2)	114	—	78 (4.8)	101	—	79 (4.8)
Ethnicity												
Caucasian and Asian	29	31		30	32		34	30		36	36	
African American	40	44		41	45		52	45		41	40	
Latino/Hispanic	23	25		21	23		27	24		24	24	
Native American	0	0		0	0		1	1		0	0	
Gender												
Male	52	57		53	58		68	60		63	62	
Female	40	43		29	32		46	40		38	38	
Special services												
Special education	3	3		2	2		3	3		1	1	
Speech therapy	6	7		3	3		8	7		2	2	
ESL	0	0		4	5		6	5		4	4	

Note. No contrasts were significantly different.

and included the previously discussed critical content. Observations of the classroom reading instruction indicated that teachers' implementation of their adopted basal programs was highly varied and that almost all teachers included other resources and methods to supplement or replace activities in the basal. Also, there was considerable variation in the classroom management styles observed, but all classroom teachers had routines and behavioral expectations that were known to the students.

Enhanced Classroom Instruction

The local district had implemented its own reading initiative for several years prior to the onset of this research, and teachers had received considerable professional development and coaching in providing comprehensive balanced literacy instruction. To build on the district's extensive professional development program, our research team worked with classroom teachers to further enhance classroom

instruction in several ways. First, we provided each teacher and principal access to several types of assessment data that were collected as part of the study. We initially shared our more comprehensive screening data, identifying for each teacher those students within the classroom who were at risk for reading difficulties and thus required greater teacher attention. For all participating students, we provided classroom teachers and intervention teachers with ongoing progress monitoring data reflective of reading growth every three weeks. These data were provided in the form of graphs of passage reading fluency for each child who participated in the study. Research has demonstrated that such ongoing progress monitoring data can enhance teacher decision making and instructional planning, improve student awareness of learning, and promote greater student achievement (e.g., Fuchs, Fuchs, Hamlett, & Stecker, 1991). Graphs of student progress in fluency of reading connected text were likewise provided

to school principals and to the parents of participating students three times during the school year.

In order to enhance teachers' use of available assessment data, we provided classroom teachers with a one-day professional development session focusing on the use of assessment data to plan and deliver differentiated instruction in the general education classroom. During the first year, this inservice focused on providing differentiated instruction based on data from the TPRI, which had been administered by first-grade teachers in each of the schools, and our one-minute passage reading fluency data. During the second year of the study, we again provided classroom teachers with a one-day inservice focused on providing differentiated instruction. In year 2, we shifted our content to focus on peer tutoring as an effective and feasible tool for accommodating academic diversity (e.g., Mathes & Fuchs, 1994; Shanahan, 1998). The final component for ensuring enhanced classroom instruction was offering the classroom teachers our services as consultants for any concerns they might have related to literacy instruction or to instructional needs of specific students. During the two years of the study, teachers often accessed the expertise of the Proactive Reading and Responsive Reading teachers who worked in their schools.

We observed the 30 teachers in the enhanced classroom condition during their language arts time three times during each year of the study, for a total of 90 observations. These observations provided information about the nature of instruction provided to students in this condition. During these observations, the observers recoded whether or not specific strategies occurred. These observers reported that the classroom was stimulating and motivating for 72.62% of the observations and that children were encouraged to express their ideas verbally in 90.27% of the observations. In terms of instructional content, instruction was provided in phonemic awareness during 34.65% of the observations, letter-sound correspondences in isolation during 72.72% of the observations, practice in reading words by sounding out

during 82.36% of the observations, and decoding words using visual memory or context in 40.53% of the observations. Comprehension strategies instruction occurred in only 15.09% of the observations, vocabulary was presented in 83.75% of the observations, spelling was taught in 70.48% of the observations, and writing was included in 54.76% of the observations. While little comprehension instruction was provided, teachers frequently assessed students' comprehension of text, asking literal questions in 83.53% of the observations and inferential questions in 61.60% of the observations.

Small-Group Intervention

Students in both supplemental intervention conditions met in small groups of three students for 40 minutes a day, five days a week, from October through May. Our decision to deliver instruction in groups of three was based on syntheses of recent research that do not identify differences in outcomes from 1:3 vs. 1:1 tutoring (Elbaum, Vaughn, Hughes, & Moody, 2000; Vaughn & Linan-Thompson, 2003). Instruction was provided at a time during the day that did not conflict with the core reading lessons offered in the regular classroom. Thus, the small-group instruction was provided in addition to enhanced classroom instruction.

The six intervention teachers who delivered this small-group instruction received 42 hours of professional development training from the authors of each intervention prior to the onset of the research. During the second year, an additional 12 hours of professional development was provided. Across both years of the study, intervention teachers also participated in monthly half-day inservice meetings. During these meetings, which were conducted separately for teachers in the two interventions, teachers (a) viewed videotaped lessons with discussion and feedback, (b) discussed issues regarding implementation of the interventions, and (c) collaborated in problem solving to plan for accelerating the growth of specific students. These teachers also received frequent onsite coaching from the intervention developers.

Proactive Reading

The objective in the design of Proactive Reading was to arrange the instructional environment to reduce the occurrence of errors and facilitate ownership and integration of skills and strategies that build cumulatively over time, and to assist students in becoming competent readers who read both fluently and with comprehension. The tasks associated with fluent, meaningful reading were analyzed and elements sequenced into a cumulatively building and carefully integrated scope and sequence. From this scope and sequence, daily lesson plans were developed. These lessons were fully specified and provided exact wording to ensure teacher language was clear and kept to a minimum. Following these prescribed lesson plans, teachers delivered explicit instruction designed to assist students in the integrated and fluent use of alphabetic knowledge and comprehension strategies. These lessons were constructed so that various content strands (i.e., phonemic awareness, word recognition, comprehension strategies) were carefully woven together.

A Typical Lesson. A primary focus of Proactive Reading was teaching efficient word identification. Thus, a large portion of each lesson was spent learning and reviewing letter–sound correspondences, sounding out and reading words rapidly, or spelling words in isolation. In a typical Proactive Reading lesson, students played word games designed to promote phonemic awareness, practiced letter–sound correspondences for previously taught letters or letter combinations, practiced writing these letters, and learned the sound of a new letter or letter combination. Students also practiced sounding out and reading words composed of previously taught letter–sound correspondences and various syllable types, spelled words from dictation based on their sound–symbol correspondences, practiced automatic recognition of words that do not conform to alphabetic rules, read and reread decodable connected text, and applied comprehension strategies to this text.

Over time, the nature of these lessons changed. In the beginning, the bulk of each

lesson was devoted to learning to use the alphabetic principle quickly and efficiently, with less focus on connected text and reading for meaning. As students progressed, lessons changed in nature to focus on decoding multisyllabic and irregular words, fluency building of connected text, and applying comprehension strategies. In later lessons, students engaged in timed readings and partner reading of narrative stories, and they engaged in retelling, with emphasis on sequencing and identifying story grammar elements.

Text Characteristics. Beginning on the seventh day of instruction, students read connected text daily. This text was fully decodable, meaning that all phonetic elements and all irregular sight words appearing in the text had been taught previously and that students had already demonstrated mastery of those elements and words. In the beginning, this text was stilted and unnatural sounding. However, as students acquired greater mastery of more and more elements, as well as the ability to decode more difficult words, this text became more and more natural.

Lesson Format. A primary feature of Proactive Reading was that it maximized academic engagement (Brophy & Good, 1986; Rosenshine & Stevens, 1986). Instruction was delivered to small homogeneous groups of students who sat in a semicircle around the instructor and was delivered in a rapid-fire manner in which there was constant interchange between the instructor and students. In a typical activity, the teacher asked all students to respond to letters, words, or text in unison, followed by individual turns when each child was able to demonstrate his or her personal ownership of the content. Moreover, the instructor moved quickly from activity to activity within each lesson. Within a typical lesson there were 7 to 10 short activities that encompassed multiple strands of content.

An overarching teaching routine repeated throughout the entire curriculum comprised the teacher modeling new content, providing guided

practice for students, and implementing independent practice for every activity. Instructors were required to consistently monitor students' responses, provide positive praise for correct responses, and provide immediate corrective feedback if an error occurred. Instructors had to make on-the-spot judgments about why an error occurred and to focus on that aspect of the task when corrective feedback was provided.

To facilitate student enthusiasm for learning, instructors provided immediate positive feedback about each activity as students demonstrated mastery. Because the curriculum was designed to gradually and cumulatively become more complex, the majority of each lesson was composed of review and generalization work. Thus, each lesson contained very little new content. The expectation was that students would enter each new activity each day able to achieve at least 80% accuracy on their first try, with 100% accuracy being achieved after error corrections and scaffolding had occurred.

Responsive Reading

Responsive Reading also provided for explicit instruction in phonemic awareness and phonemic decoding, but it dedicated relatively less time to the practice of these skills in isolation than did the Proactive approach. Students in the Responsive intervention spent about one fourth of their daily 40-minute lesson in isolated skills instruction and practice. During the balance of the lesson, students in Responsive Reading were prompted to apply literacy skills and strategies in the context of extensive reading and writing practice.

A Typical Lesson. In contrast to Proactive Reading, Responsive Reading did not include a predetermined scope and sequence. Responsive teachers used data from student assessments and daily anecdotal records to identify student needs and strengths, and they planned their instruction based on this analysis, attempting to provide instruction and support within students' zone of proximal development (Vygotsky, 1978)—the level of difficulty

at which students could be successful with the support of the teacher. Teachers individualized instruction by focusing their daily lesson planning and text selection on an individual student within each group, alternating between the three students each day. Thus, every three days, each child in the group received this concentrated attention from the teacher. Each day, the teacher sat beside the focus child and directed instructional scaffolding, prompting predominantly to that child while including the others in the lesson activities.

Responsive Reading teachers adhered to a lesson cycle that demarcated how time was used across each 40-minute lesson. This cycle had five components: fluency building, assessment, letter and word work, supported reading, and supported writing. Teachers were required to choose activities from a menu of options for each part of the lessons based on the observed needs of their students. The nature of these activities, as well as the texts the students read in the lessons, became more complex over time.

During the first two components, fluency building and assessment, which lasted 8 to 10 minutes, students engaged in repeated reading with teacher modeling and prompting to support passage reading fluency. The teacher modeled fluent and expressive reading, explicitly taught the meaning and oral interpretation of punctuation marks, and prompted students to read smoothly and in phrases. The instruction and modeling were directed to the daily focus student, while the other two students in the group engaged in partner reading. Following the fluency activities, the teacher observed one student reading a book that had been introduced on the previous day, using running record procedures (see Clay, 2002). The teacher individually assessed the reading strengths and needs of each student in the group one to two times per week.

During the third component, letter and word work, students received 10 to 12 minutes of explicit instruction and practice related to phonemic awareness, letter-sound relationships, word reading, or spelling. Teachers explicitly taught phonological awareness skills, letter-sound

correspondences, and how to sound out words, and students reviewed and practiced these skills daily. Students were taught to manipulate onsets and rimes to arrive at new words through analogy with known words, but they were also asked to segment the phonemes within each onset and rime before applying these larger units to read words. Students practiced decoding and encoding skills in several formats, including blending and sounding out words presented by the teacher and segmenting and writing dictated words.

The fourth component, supported reading, lasted 10 to 12 minutes. During this time, students read a text they had not previously read. Each day, the focus student read a portion of the book alone, while the teacher coached, scaffolded, and prompted the student to apply skills and strategies. Next the entire group read the same book, either chorally or individually. Prior to the reading of a new book, teachers pretaught potentially difficult vocabulary words, discussed potentially confusing subject matter, and encouraged students to make predictions to link the book's subject matter to prior knowledge and to establish a purpose for reading. During and after reading, teachers frequently asked questions referring to the text meaning and asked students to retell or summarize portions of the story. Teachers provided feedback and supported students as they discussed the meaning of text.

The final lesson component, supported writing, consisted of 8 to 10 minutes of sentence writing with the teacher providing coaching and scaffolding in the application of the strategy of sound analysis for the spelling of unknown words. During supported writing, teachers assisted students in writing sentences about the new story in journals. Teachers sometimes provided explicit instruction in word patterns, modeled the segmenting of words in order to record their phonemes, or used Elkonin sound boxes (Elkonin, 1973) as a framework to assist students in the application of the alphabetic principle.

Word Identification Strategies. As students practiced reading words in isolation

and reading connected text, there was a clear emphasis on teaching efficient word recognition. The following was the primary word recognition strategy taught in Responsive Reading: (a) look for parts you know (i.e., identify known letter combinations), (b) say the word slowly and blend the sounds (i.e., sound it out), and (c) reread the sentence with the word in it and decide whether it makes sense (i.e., check the word within context). Students were also taught to decode unknown words using analogy to known words. Students were taught to access semantic and syntactic information primarily to self-monitor their reading. Contextual information was used to support decoding only occasionally, always in conjunction with the primary strategy of phonemic decoding. Likewise, students were taught to segment words and apply letter-sound associations as a strategy for spelling unknown words.

Text Characteristics. In our study, the Responsive teachers selected books ranked for difficulty in 16 different levels for use in guided reading instruction (Fountas & Pinnell, 1999). These books were leveled for difficulty but were not intended to be phonetically decodable. They progress in complexity of word types and syntax as well as aspects of print size and page layout (see Peterson, 1991).

Comparison of the Two Interventions

Key distinctions between the Proactive and Responsive interventions can be categorized into four main differences. First, Proactive Reading provided a detailed scope and sequence with fully specified daily lessons. In contrast, Responsive Reading relied on the teacher to plan instruction based on observed student needs and strengths. Second, the type of text students read in each intervention was different. Proactive Reading exclusively relied on decodable text, which in the beginning was not particularly natural sounding but became more natural and engaging over time. Responsive Reading made use of leveled text that was not phonetically decodable, but was

more natural sounding and arguably more engaging even from the beginning. Third, these two interventions differed in the amount of time focused on learning skills in isolation versus in context. Students in Proactive Reading spent substantial time practicing skills and words in isolation, contrasted with the time Responsive Reading students spent applying strategies and skills with teacher support while they engaged in reading and writing connected text. Finally, the interventions differed in the incorporation of a writing component. Writing in Proactive Reading was limited primarily to spelling words in isolation, while students in Responsive Reading spent about nine minutes of each lesson learning to record their thoughts in complete sentences, while also receiving instruction in the application of the alphabetic principle in spelling words.

Intervention Fidelity

In order to ensure that each intervention was conducted as described, the first and second authors observed intervention teachers and students during an entire instructional session every eight weeks for a total of four observations of each teacher. A 3-point rating scale was used to evaluate the fidelity of implementation of each activity or section of a lesson across four categories: (a) appropriate pacing, (b) implementation of prescribed procedures, (c) error correction with appropriate scaffolding, and (d) student engagement and attentiveness. A score of 3 indicated that the teacher implemented the category in exactly the way it was intended. A score of 2 indicated that the category was implemented in an acceptable manner, but with some error. A score of 1 indicated the category was poorly represented. Likewise, we included a global checklist for readiness of instructional materials, appropriate student seating arrangement, and instructor warmth and enthusiasm. On average, both sets of intervention teachers conducted their respective interventions with high levels of fidelity.

Proactive teachers were rated as having materials ready, having students seated appropriately, and having a warm and enthusiastic

manner 100% of the time. Likewise, they were rated as having good instructional pacing ($M = 2.75$, $SD = .53$), following procedures by presenting the lesson as prescribed ($M = 2.64$, $SD = .65$), correcting errors and scaffolding appropriately ($M = 2.71$, $SD = .57$), and maintaining student attentiveness ($M = 2.84$, $SD = .44$).

In a similar manner, Responsive teachers were rated as having materials ready 96% of the time, having students seated appropriately 96% of the time, and being warm and enthusiastic 100% of the time. They were rated as having appropriate pacing ($M = 2.48$, $SD = .73$), following procedures for each section of the lesson routine ($M = 2.59$, $SD = .62$), scaffolding student responses appropriately ($M = 2.67$, $SD = .54$), and maintaining student attentiveness ($M = 2.74$, $SD = .55$). No statistically significant differences were detected in levels of fidelity between the Proactive and Responsive groups, with the exception of pacing, in which the Proactive intervention was rated slightly, but reliably better, $t(252) = 3.69$, $p < .001$.

Measures

Rationale. Proficiency in reading requires, at a minimum, that children be able to read words and text accurately and fluently and understand the meaning of the text. These outcomes are consistent with the content of the reading interventions. Thus, to address the first research question, which concerned the effectiveness of the reading interventions, the measures included assessments of word reading, fluency, and reading comprehension. Given the differences in the amount of time devoted to these different reading domains, we predicted that the Proactive intervention might have more impact on word reading, while Responsive might have more impact on fluency and comprehension. Assessments were done throughout the year to assess the impact of the interventions on growth in word reading and fluency. Assessments of word reading, fluency, and comprehension using norm-referenced tests were conducted at the end of the school year to determine the effects of the interventions on

practically important outcomes. This approach permits multiple measures of these constructs as a way of increasing reliability for the assessment outcomes, which is important in ensuring that the results are not dependent on how the outcome was measured. In addition, we did not want the results to be dependent on a particular form of assessment. Thus, fluency was assessed for words and text, and comprehension was assessed with a cloze procedure in which the child filled in missing words from a passage and answered comprehension questions about longer passages. Because the two interventions addressed word recognition instruction differently, separate assessments of accuracy and fluency of real word reading and pseudoword reading were conducted. Phonological awareness was also assessed during the year as both programs explicitly taught phonological awareness. Our approach to assessment of outcomes for the first research question was thus consistent with reading theory and with the instructional content of the two programs.

Some of these measures, in combination with a measure of rapid serial naming of letters, could also be used to address the second research question, whether specific child characteristics predicted response to different types of intervention. Phonological awareness is most strongly linked to word recognition, rapid naming to fluency, and vocabulary to comprehension. Thus, we predicted that children who were weaker in phonological awareness might make more progress with the Proactive intervention, while children weaker in speed of phonological retrieval and vocabulary might do better with Responsive.

Description of Growth Assessments. The assessments of growth were administered four times during the school year at two-month intervals beginning in October, and they included phonological awareness, rapid automatized naming of letters (RAN letters), untimed word reading, word reading fluency, nonword reading fluency, and passage reading fluency (done every three weeks). Phonological awareness was measured using the First Sound

Comparison, Blending Onset-Rime, Blending Words, Blending Nonwords, and Phoneme Elision subtests from the Comprehensive Test of Phonological Processes (CTOPP). Children's scores on the various phonological awareness measures were placed on a single scale based on item response theory (IRT) work (Schatschneider et al., 1999) so that individuals could be assigned a single score of overall phonological awareness ability. This phonological awareness score, reported as a theta score, is much more reliable than individual subtest scores or a typical composite score and is thereby more sensitive and more appropriate for monitoring growth. The final subtest from the CTOPP administered in this research was the RAN Letters subtest, which measures a child's efficiency of phonological access. In this subtest, students are timed as they provide the names of a set of known letters for 60 seconds. The range of reliabilities (internal consistency) on the CTOPP subtests for this age group range from .70 to .91, all acceptable or better.

Untimed word reading entailed reading words from an IRT-based list of increasingly more difficult words developed by our research team. This list has been found to be very sensitive to short-term growth in word recognition ability (see Foorman et al., 1998). The words on the list were selected according to frequency and diversity of linguistic and orthographic features represented in early primary texts. The internal consistency of this measure is .90 (O'Malley, Francis, Foorman, Fletcher, & Swank, 2002).

Word reading fluency and nonword reading fluency were measured using the Sight Word Efficiency and Phonemic Decoding Efficiency subtests from the Test of Word Reading Efficiency (TOWRE). In these subtests, students read as many words as they could in 45 seconds or decoded as many pseudowords as they could in 45 seconds. Each list of words and nonwords was arranged so that items increased in difficulty. We included both words and nonwords to ensure that we measured both phonological decoding ability and sight recognition

of familiar or partially familiar words. Internal consistency exceeds .95 for both subtests.

Passage reading fluency was measured as words read correctly per minute (WCPM) on timed one-minute oral reading samples of end-of-first-grade-level passages that had been developed for Teachers Continuous Assessment for Reading Excellence software (TCARE; Mathes, Torgesen, & Heron, 2004). The passages used to evaluate oral reading fluency were subjected to substantial field-testing to determine equivalence of difficulty.

Description of End-of-Year Assessments.

Measures that were administered only at the end of the school year (i.e., posttest only) included the WJ-III, the Comprehensive Assessment of Reading Battery Revised for First-Grade (CRAB-R; see Mathes et al., 1998), and the Vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (WASI). From the WJ-III, the Word Attack, Word Identification, Passage Comprehension, Reading Fluency, Spelling, and Calculations subtests were administered. Reliability ranges from .87 to .97. The Word Attack subtest is a measure of accurate decoding of nonwords, whereas Word Identification is a measure of the ability to read sight words in lists. Reading Fluency measures how quickly students read sentences and determine if they are true or false. Passage Comprehension is measured through a cloze procedure, where students read a sentence or brief passage in which certain words have been taken out and students are required to produce the missing words or acceptable substitutions for them. Spelling requires students to spell words of increasing complexity. The test of Mathematical Calculation was included to determine whether students made generalized academic progress or gains specific to reading acquisition. The CRAB-R yields fluency and comprehension scores with test–retest reliability of .92 to .96 for different indexes. For this test, students read two stories orally for three minutes each and answered 10 questions about each story. The CRAB-R was not administered

to students who had a raw score of 4 or less on WJ-III Passage Comprehension.

Results

Analytic Approach

Multilevel modeling techniques represented the primary data analysis procedures. Research on students in schools represented nested levels of analysis because each student was embedded in a classroom, which in turn was embedded in a school. Although our research questions focused on the student level of analysis, outcomes were affected by classroom-level differences (e.g., curriculum, classroom teacher styles) and school-level differences (e.g., variations in the number of high-poverty students at a school). If growth was assessed, this represented another level in the design (i.e., time) that was nested within students. Thus, multilevel modeling permitted us to address the primary research questions while controlling for sources of variability that could produce potentially unexplained sources of differences in the outcomes. When time was included in the analysis, as was the case for the growth measures, this approach allowed us to estimate growth curves for each student relative to the entire samples' growth pattern. With growth curves, we could (a) estimate the average rate of change for the entire sample, represented by the slope parameter; (b) estimate the extent to which individual growth rates differed from this average rate of change; (c) estimate the average level of performance at the final time point in April for the entire sample, represented by the intercept parameter; (d) estimate the extent to which individuals differed from this average April performance; and (e) evaluate correlates of change while controlling for classroom effects. The same multilevel modeling approach was used with the norm-referenced, end-of-year data, except that time was not included in the model and rates of growth were therefore not computed. Analyses that addressed hypotheses 1, 2, and 3 focused on the intervention group, as it predicted the intercept or slope parameters, as the correlate

of change. Analyses that addressed hypothesis 4 examined student characteristics as correlates of change that may have interacted with intervention.

Prior to predicting individual students' growth patterns, it was necessary to accurately characterize the overall growth pattern of the sample for each outcome. Accurate descriptions of the sample's general growth patterns were achieved through a process called building the *unconditional models*. Growth can be linear, characterized by a slope that is a straight line. It can also be nonlinear, which is common in grade 1 (see Foorman et al., 1998). Thus, it was necessary to test models with linear and curvilinear growth.

Building accurate unconditional models of the sample's growth patterns also required that we examine whether there was significant variability among individual children's growth patterns. Examination of individuals' variability from the average growth pattern involved testing whether the change parameters (e.g., intercept and slope) should be fixed or freely estimated. Fixed parameters have the same values for all children. Random parameters are freely estimated across children and quantify the degree of individual variability.

Characterizing the present sample's growth patterns involved sequentially testing a series of models that increased in complexity by one change parameter per model to determine which model best characterized the sample's growth. The order of the five models that could have potentially been examined were as follows: (a) straight line growth with random intercept and fixed slope; (b) straight line growth with random intercept and random slope; (c) curvilinear growth with random intercept, fixed slope, and fixed quadratic; (d) curvilinear growth with random intercept, random slope, and fixed quadratic; and (e) curvilinear growth with random intercept, random slope, and random quadratic. If the mean value of the target change parameter in a given model was significantly different from 0 using a criterion alpha level of .05, then that parameter was kept and model building continued in the order above. If

the new parameter being tested was not significant, then that parameter was dropped from the model and model building ceased.

Identifying parameters that should or should not be in the unconditional model is not the same as testing for significant differences between groups, even though a *t* statistic is reported for both. It is possible that the number of tests of significance of parameters could lead to Type I errors in some instances, falsely identifying a parameter as important, but such errors are less critical than failure to identify a potentially significant parameter. Setting alpha at .05 provided an appropriate balance for Type I and II errors. We centered growth trajectories at the final wave of assessment so that intercepts represented expected levels of achievement in April. This approach is standard practice for growth curve modeling and is necessary to develop unconditional models that accurately, parsimoniously, and understandably characterize change within the sample.

Once the best unconditional models were determined, change parameters were correlated with relevant predictors of growth, such as intervention group or children's initial status on reading-related skills. These *conditional models* directly tested the effectiveness of the interventions and whether specific child characteristics differentially predicted change associated with an intervention. We tested each outcome separately because a univariate approach is consistent with our interest in replicating findings across measures of the same construct that vary in assessment method, thus avoiding method dependence of the findings.

Magnitude and Reliability of Effects

Hypotheses 1, 2, and 3 involved follow-up comparisons of intervention-group effects. Hypothesis 4 involved a few follow-up comparisons of intervention group by child characteristic interaction effects. All of these comparisons were conducted using a critical alpha level of .05. A Bonferroni adjusted alpha of .01 was not used because it was judged too conservative and would inflate the Type II error rate. Instead, our use of a .05 criterion

was consistent with our goal of using the current sample and being able to detect practically important and educationally significant effects of at least moderate effect size. The tables indicate statistical significance at multiple criterion alpha levels so that readers can independently judge the contrasts if they desire to do so. The second way we evaluated our interventions did not rely on an arbitrarily determined criterion but instead involved examining effect sizes. This helped us quantify the magnitude of group differences on the intercept (i.e., scores in April or May) and slope (i.e., growth rates) of the two supplemental intervention groups relative to the intercept and slope of the enhanced classroom group. Effect sizes were calculated for the intercept and slope terms by subtracting the estimates of the enhanced classroom group from the estimates of the intervention groups and dividing by the square root of the residual.

Analyses to Identify Confounding Variables

Prior to addressing our two primary research questions, it was necessary to consider the influence of differences among the groups that might influence the outcomes. First we examined differences in initial status. Differences in initial levels of literacy development between the typically achieving and at-risk groups were expected. However, we did not expect the three at-risk groups to differ in these characteristics. Second, we examined other student-level variables that would represent potential covariates in the analyses. These included gender, ethnicity, classroom, and cohort. Cohort had to be considered because we conducted the study over two consecutive years to increase the sample size. Covariate effects are reported for each analysis in which they were significant. There were no instances where any covariate interacted with intervention.

Analyses of Baseline Differences in October

Multilevel modeling was used to control for classroom effects and test for group differences

in initial status on the following baseline measures: TPRI letter names and sounds (used to identify students' risk status), untimed word reading, TOWRE word reading fluency, TOWRE nonword reading fluency, TCARE passage reading fluency, CTOPP phonological awareness IRT-based theta score, and WASI Vocabulary subtest. There were significant overall group differences on all of these baseline variables, $F_s(3, 330-356) = 8.2$ to 94.7 , $ps < .001$. As expected, the typically achieving group had higher language and literacy scores than the three at-risk groups. Generally, the three at-risk groups had comparable baseline language and literacy scores. The only exception was that students in the Responsive group had higher baseline scores on letter names than students in the enhanced classroom group. Baseline scores for each group are presented in Table 2.

Hypotheses 1, 2, and 3: Effectiveness of the Interventions

We examined effectiveness of the interventions relative to quality classroom instruction for at-risk students, relative to quality classroom instruction for typically achieving students, and relative to each other in terms of rate of growth (slope) and April levels of performance (intercept) for each growth measure. These three contrasts were also used to examine the effectiveness of the interventions on end-of-year (May) outcomes.

Analyses of Growth

Table 3 reports comparisons of the slope, intercept, and, when appropriate, quadratic values for all growth measures among the four groups.

Phonological Awareness. The sample's growth in phonological awareness was best described by a model that included random intercept, random slope, and random quadratic terms. In other words, the unconditional model

Table 2
Baseline Scores by Group

Measure	Observed Scores by Group							
	Enhanced Classroom (n = 114)		Proactive (n = 92)		Responsive (n = 92)		Typically Achieving (n = 101)	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Word reading ^a	-1.25	(.65)	-1.30	(.64)	-1.22	(.61)	.29	(1.1)
CTOPP								
Phonological Awareness ^a	-.77	(.62)	-.87	(.67)	-.79	(.64)	-.02	(.57)
RAN Letters ^b	.84	(.28)	.81	(.25)	.80	(.22)	1.13	(.25)
Nonword Repetition ^c	9.5	(4.2)	9.2	(3.8)	9.5	(3.5)	11.6	(4.2)
TPRI								
Letter Names ^c	9.2	(1.4)	9.2	(1.5)	9.5	(1.0)	9.9	(.50)
Letter Sounds ^c	6.1	(2.9)	6.1	(3.1)	6.4	(2.7)	9.0	(1.3)
TOWRE								
Word Reading Efficiency ^d	86.2	(8.3)	86.2	(9.0)	86.8	(7.4)	95.9	(12.2)
Nonword Reading Efficiency ^d	93.0	(8.2)	93.4	(7.7)	93.1	(6.7)	100.8	(10.4)

^aIRT-based theta scores; ^bLetters per minute; ^cRaw scores; ^dStandard scores.

indicated that there were reliable individual differences in the rates of growth, the degree of curvilinearity in that growth, and April scores. The unconditional model also included a random intercept term at the classroom level that modeled reliable classroom differences on April phonological awareness scores. Examination of potential covariates found significant effects of ethnicity only on individuals' expected phonological scores in April, $F(2, 50) = 11.27, p < .001$. Subsequent analyses of growth in phonological awareness controlled for effects of ethnicity and classroom.

The overall impact of intervention group was significant on the intercept, $F(3, 85) = 15.71, p < .001$, and slope $F(3, 1431) = 29.25, p < .001$. Follow-up comparisons of the groups' slope values (see first column of Table 3) indicated that the Proactive and Responsive groups demonstrated significantly more rapid development of phonological awareness than the enhanced classroom and typically achieving

groups. In addition, the Proactive group demonstrated more rapid growth than the Responsive group. Growth patterns for each group are illustrated in Figure 1.

Follow-up comparisons of the groups' intercept values, which represent predicted April scores, indicated that the Proactive and Responsive intervention groups had higher April scores than the enhanced classroom group (see Table 4 and Figure 1). However, the two intervention groups continued to have April scores that were lower than those of the typically achieving group. Finally, the Proactive and Responsive groups had comparable phonological awareness scores in April.

Untimed Word Reading. The sample's growth in untimed word reading ability was best described by an unconditional model including random intercept, random slope, and fixed quadratic terms, indicating that there were significant amounts of individual variation in

Table 3
Group Differences in Growth Patterns of Literacy Skills

Contrasts	Phonological Awareness			Untimed Word Reading			Word Reading Fluency			Nonword Reading Fluency			Passage Reading Fluency (Cohort 1)			Passage Reading Fluency (Cohort 2)									
	diff.	(SE)	t	diff.	(SE)	t	diff.	(SE)	t	diff.	(SE)	t	diff.	(SE)	t	diff.	(SE)	t							
Slope																									
P vs. EC	0.16	(.02)	1431	6.72***	0.17	(.04)	1428	4.26***	3.90	(1.18)	1428	3.28**	0.95	(.41)	1436	2.30	0.27	(.53)	1935	0.51	2.23	(.80)	2370	2.77***	
R vs. EC	0.08	(.02)	1431	3.22**	0.11	(.04)	1428	2.73**	2.06	(1.18)	1428	1.75+	0.34	(.41)	1436	0.83	0.77	(.53)	1935	1.47	2.04	(.79)	2370	2.57*	
P vs. R	0.08	(.02)	1431	3.40***	0.06	(.04)	1428	1.49	1.83	(1.21)	1428	1.51	0.61	(.41)	1436	1.44	0.50	(.55)	1935	0.92	0.19	(.82)	2370	0.23	
P vs. TA	0.21	(.02)	1431	8.76***	0.46	(.04)	1428	11.50***	3.55	(1.18)	1428	3.01**	0.88	(.41)	1436	2.11*	1.83	(.51)	1935	-2.97**	2.26	(.80)	2370	2.82***	
R vs. TA	0.12	(.02)	1431	5.27***	0.39	(.04)	1428	10.01***	1.71	(1.17)	1428	1.47	0.27	(.41)	1436	0.65	1.56	(.53)	1935	-2.06*	2.08	(.79)	2370	2.63***	
Intercept																									
P vs. EC	0.34	(.08)	85	4.34***	0.37	(.08)	85	4.75***	4.57	(2.10)	85	2.17*	2.43	(1.37)	85	-1.78	1.12	(5.59)	56	-0.20	10.91	(5.54)	53	1.97+	
R vs. EC	0.19	(.08)	85	2.45*	0.31	(.08)	85	4.03***	4.83	(2.08)	85	2.33*	1.08	(1.36)	85	-0.80	7.79	(5.56)	56	-1.40	12.23	(5.52)	53	2.22*	
P vs. R	0.15	(.08)	85	1.84+	0.05	(.08)	85	0.74	0.26	(2.16)	85	0.12	1.36	(1.41)	85	0.95	6.67	(5.79)	56	-1.15	1.32	(5.70)	53	0.23	
P vs. TA	-0.16	(.08)	85	-2.05	-0.18	(.08)	85	2.38*	-15.50	(2.13)	85	-7.28***	-6.58	(1.39)	85	-4.74***	-30.35	(5.57)	56	-5.45***	-25.85	(5.85)	53	-4.63***	
R vs. TA	-0.31	(.08)	85	-3.97***	-0.24	(.08)	85	-3.17**	-15.24	(2.10)	85	-7.27***	-7.93	(1.34)	85	-5.81***	-23.68	(5.43)	56	-4.36*	-24.54	(5.56)	53	-4.41***	
Quadratic																									
P vs. EC					0.75	(.30)	1428	2.54*													0.11	(.05)	2370	2.40*	
R vs. EC					0.29	(.30)	1428	0.96														0.81	(.05)	2370	1.82+
P vs. R					0.47	(.31)	1428	1.52														0.03	(.05)	2370	0.59
P vs. TA					0.89	(.30)	1428	3.00**														0.26	(.05)	2370	-5.80
R vs. TA					0.42	(.30)	1428	1.41														0.23	(.05)	2370	5.26***

Note. P = Proactive; R = Responsive; EC = enhanced classroom; TA = typically achieving; diff = difference between the estimates; df = degrees of freedom.
 + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

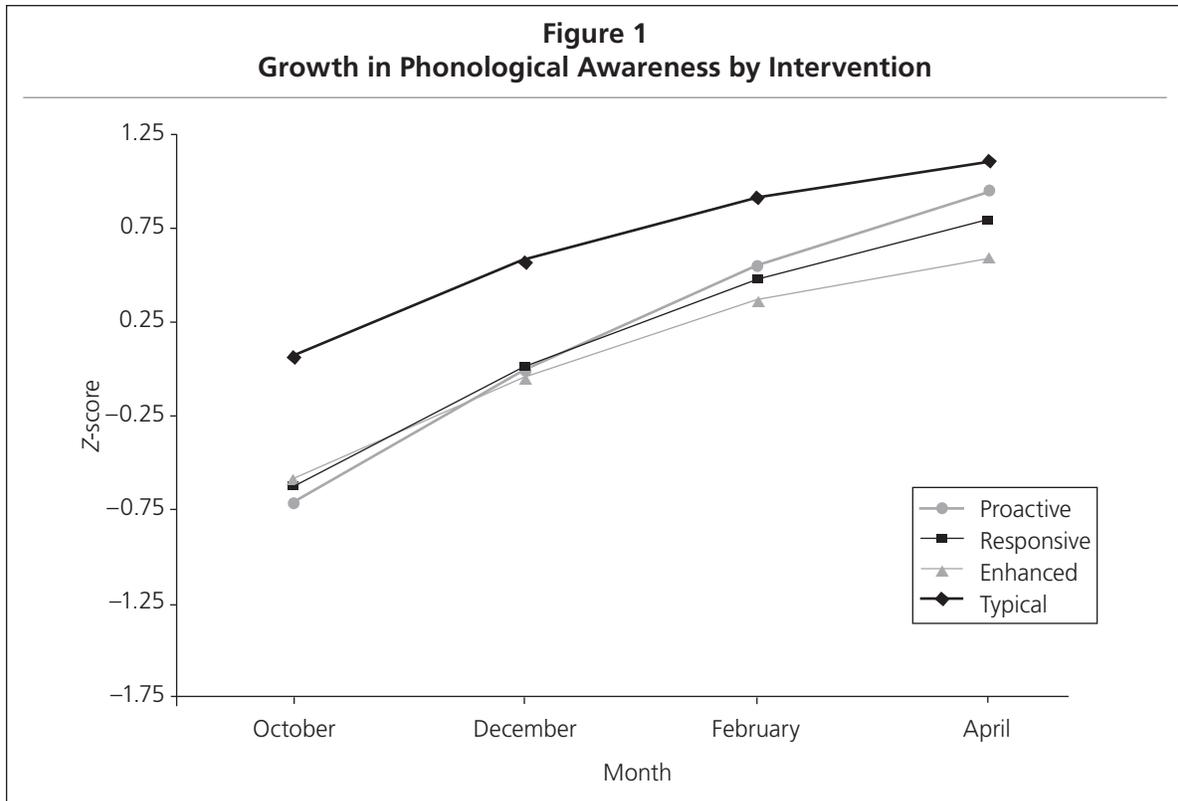


Table 4
Intervention Effect Sizes Compared to Enhanced Classroom Instruction on Slope and Intercept Terms

Measure	Proactive		Responsive	
	Slope	Intercept	Slope	Intercept
Phonological awareness	.81	1.76	.39	.99
Untimed word reading	.47	1.03	.30	.87
TOWRE				
Nonword reading fluency	.21	.52	.07	.23
Word reading fluency	1.13	1.33	.60	1.41
Passage reading fluency				
Cohort 1	.04	.15	.11	1.06
Cohort 2	.33	1.62	.30	1.81

Note. Slope represents growth and intercept represents April scores.

rates of growth and April scores. The unconditional model also included a random intercept term at the classroom level. This parameter indicated there were reliable classroom differences in April scores. There were no effects of

cohort, gender, or ethnicity that needed to be covaried.

The overall effect of group was significant on the slope, $F(3, 1428) = 53.22, p < .001$, and intercept, $F(3, 85) = 18.70, p < .001$. Follow-up

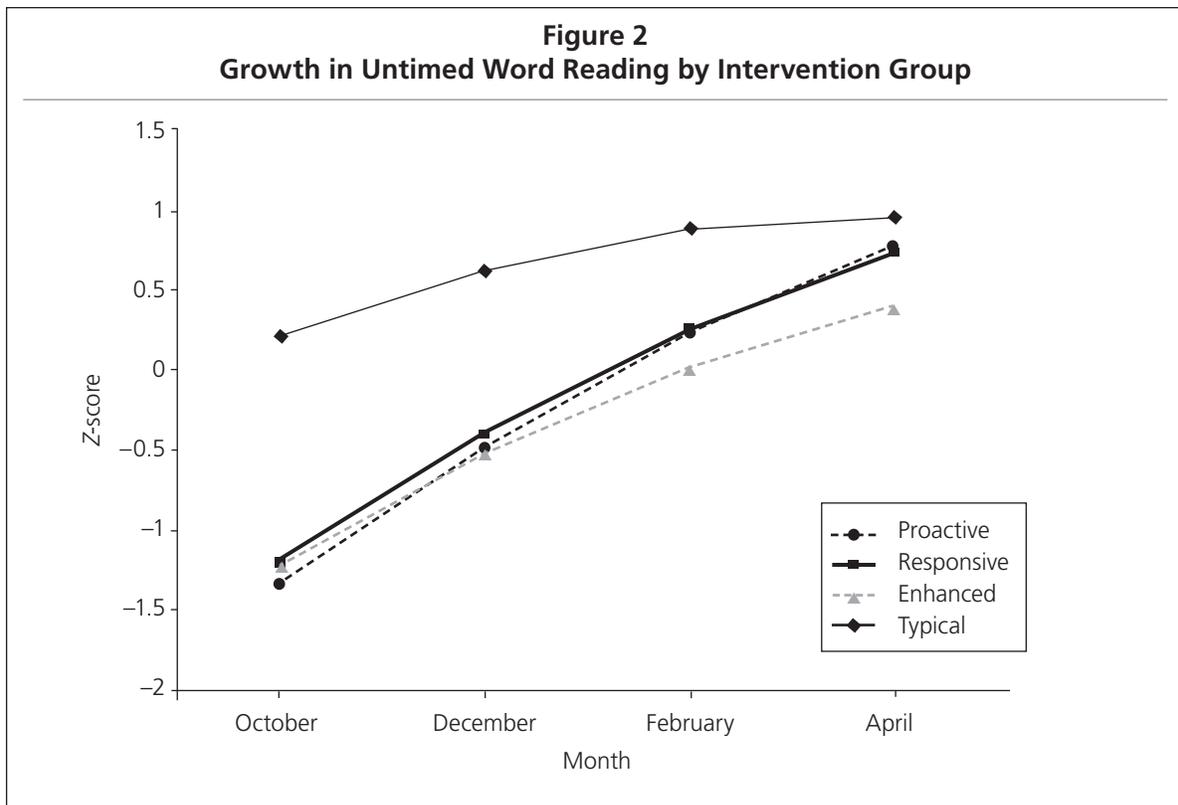
comparisons of the groups' slopes indicated that both the Proactive and Responsive intervention groups demonstrated significantly more rapid development than that of the enhanced classroom and typically achieving groups (see second column of Table 3 and Figure 2). Finally, rates of growth in untimed word reading among the two intervention groups were comparable.

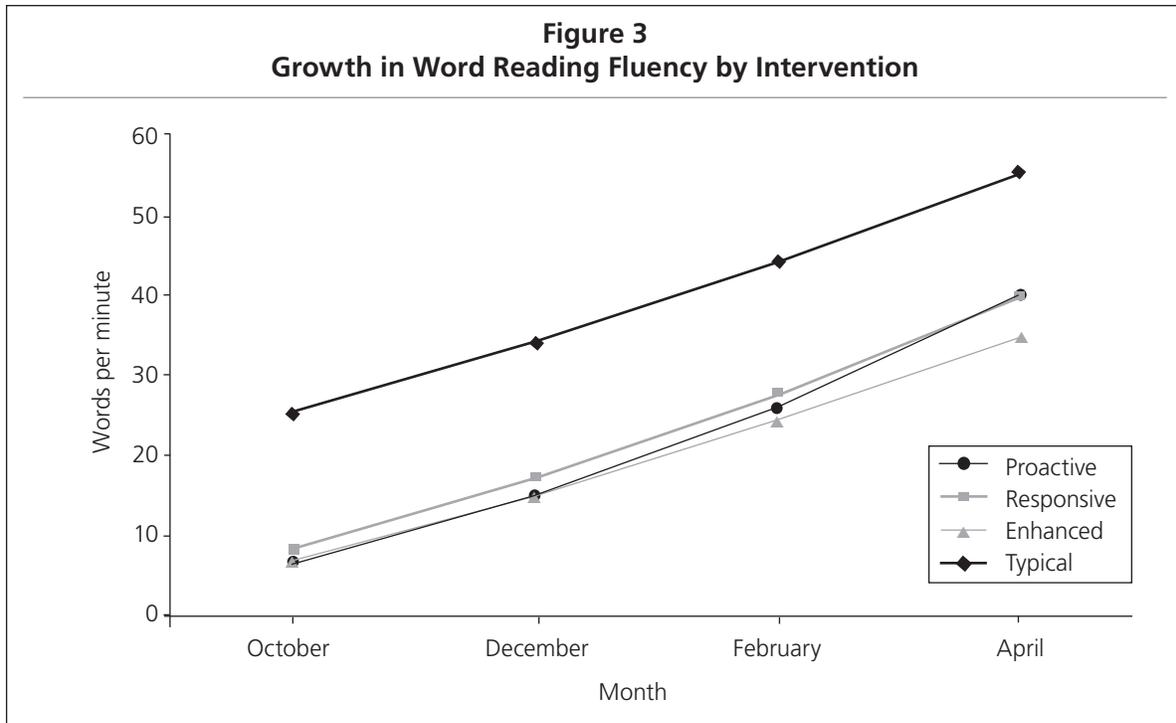
Comparisons of the groups' intercepts revealed that both intervention groups had April scores in untimed word reading that were significantly higher than those of the enhanced classroom group (see Table 3). Both intervention groups had April scores that continued to be lower than those of the typically achieving group. Finally, the two intervention groups had comparable scores on untimed word reading in April.

Word Reading Fluency. The sample's growth in TOWRE Word Reading Fluency was best

described by an unconditional model with random intercept, random slope, and random quadratic terms, indicating there were reliable individual differences in rates of growth, curvilinearity of growth patterns, and April scores. The unconditional model also included random intercept and random slope terms at the classroom level. These later parameters modeled reliable classroom differences in rates of growth and April scores. Among potential covariates, there was a small but statistically significant effect of cohort on individuals' degree of curvilinearity of growth in word reading fluency, $F(1, 437) = 4.33, p < .05$, such that the second cohort had slower initial growth but similar April scores.

There were significant differences in growth patterns in word reading fluency between the four groups, controlling for classroom and cohort, as illustrated in Figure 3. The overall effect of intervention group was significant on





the intercept, $F(3, 85) = 36.08, p < .001$; slope, $F(3, 1428) = 4.53, p < .01$; and quadratic terms, $F(3, 1428) = 3.46, p < .05$. The Proactive group demonstrated more rapid development than the typically achieving and enhanced classroom groups (see comparisons of slope terms in third column of Table 3). The Proactive group also showed significantly more acceleration in their rate of growth relative to the typically achieving and enhanced classroom groups (see comparisons of quadratic terms in Table 3 or increasingly steeper slope of the Proactive group in Figure 3). The Responsive group failed to demonstrate a growth rate or acceleration pattern that was significantly different from the typically achieving or enhanced classroom groups. Finally, there were no statistically significant differences between the growth rates and acceleration patterns of the two intervention groups.

Comparisons of the intercept values found that both of the intervention groups had word reading fluency scores in April that were higher than those of the enhanced classroom group.

The intervention groups continued to have lower April scores than the typically achieving group, and they were comparable to one another.

Nonword Reading Fluency. The sample's growth in TOWRE nonword reading fluency was best described by a model including random intercept and random slope terms, indicating there were reliable individual differences in linear rates of growth and April scores. The unconditional model also included random intercept and random slope terms at the classroom level, indicating there were reliable classroom differences in rates of growth and April scores. There were no effects of cohort, gender, or ethnicity. There were significant group differences in students' growth in nonword reading fluency, controlling for classroom. The overall effect of group was significant on the intercept, $F(3, 85) = 18.11, p < .001$, but not the slope, $F(3, 1436) = 2.14, p = .09$ (see Table 3). Only the Proactive group demonstrated a statistically significant quicker growth

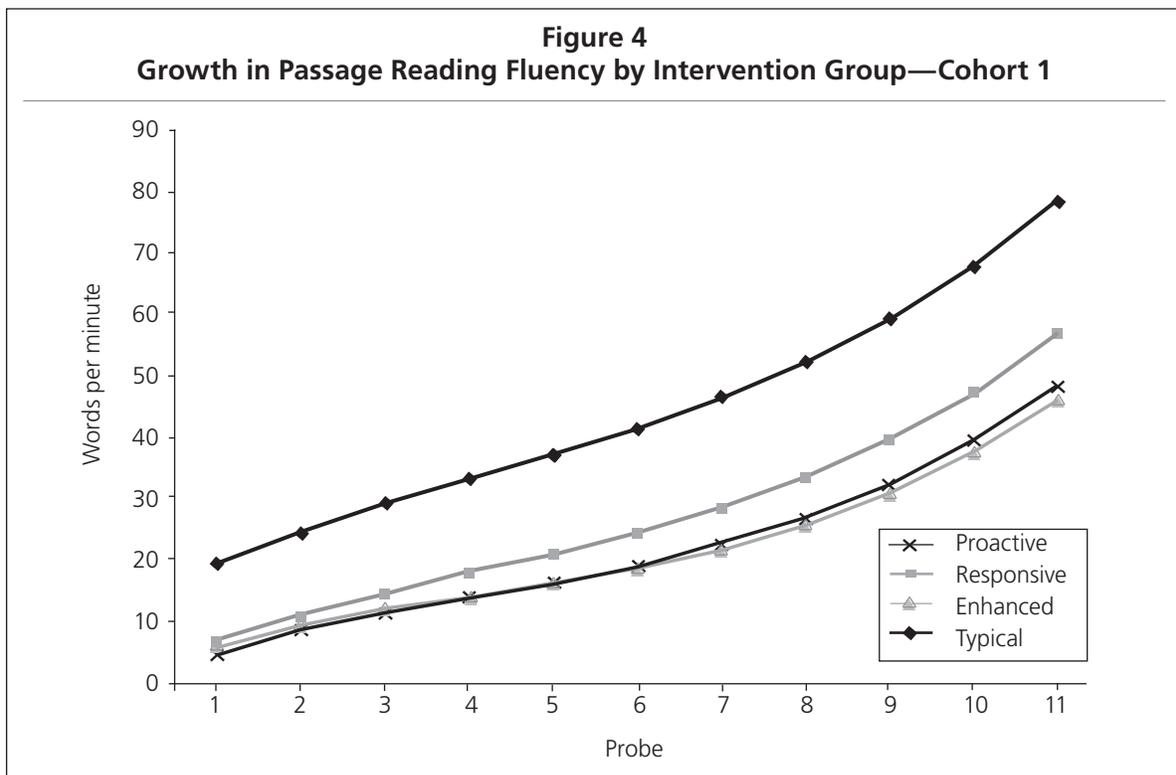
rate than the enhanced classroom group or the typically achieving group. The intervention groups had comparable growth rates. All three at-risk groups had comparable April scores on nonword reading fluency, and all three at-risk groups had lower April scores than the typically achieving group.

Passage Reading Fluency. We examined cohorts separately on passage reading fluency because we collected a different number of one-minute oral reading samples in year 1 and year 2 and because the same stories were administered at different time points during those years.

Cohort 1. The unconditional model that best described the sample’s improved ability to fluently read connected text during the first year of the study was a model with random intercept, random slope, random quadratic, and fixed cubic terms. In other words, there were reliable

individual differences in rates of growth, curvilinearity of growth patterns, and April scores. The unconditional model also included random intercept and random slope terms at the classroom level, reflecting reliable classroom differences in growth rates and April passage reading fluency scores.

There were significant group differences in growth of passage reading fluency, controlling for classroom (see Figure 4). The overall effect of group was significant on the intercept, $F(3, 56) = 14.58, p < .001$, and slope, $F(3, 1935) = 5.03, p < .01$. The typically achieving group demonstrated significantly more rapid development (slope) than the Proactive, Responsive, and enhanced classroom groups. Both intervention groups demonstrated growth rates comparable to those of the enhanced classroom group and to each other. Accordingly, the typically achieving group had higher April scores (intercept) than all three at-risk groups, and the three at-risk groups had comparable April scores.

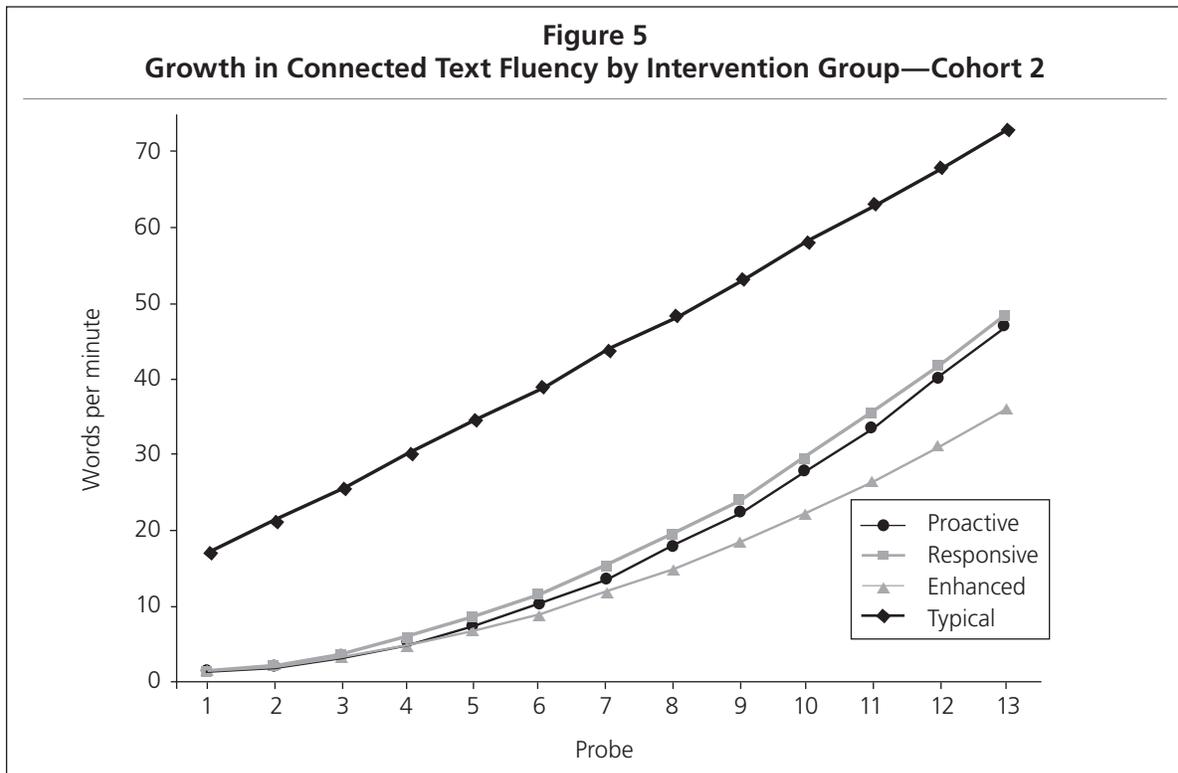


Cohort 2. Growth in ability to read connected text in year 2 was best described by a model including random intercept, random slope, and random quadratic terms. There were reliable individual differences in rates of growth, curvilinearity of growth patterns, and April scores. The unconditional model also included a random intercept term at the classroom level, reflecting reliable classroom differences in April passage reading fluency scores. Among potential covariates, there was a significant effect of gender on the intercept, $F(1, 17) = 9.94, p < .01$; slope, $F(1, 2379) = 6.64, p = .01$; and quadratic, $F(1, 2379) = 3.57, p = .05$. Specifically, the passage reading fluency of girls started off higher and increased more rapidly than for boys.

There were significant group differences in growth in reading connected text, controlling for classroom and gender. The overall effect of group was significant on the intercept, $F(3, 53) = 16.45, p < .001$; slope; and quadratic terms,

$F(3, 2370) = 4.86$ and $13.98, ps < .01$, respectively. The three at-risk groups demonstrated more accelerated rates of growth relative to the typically achieving group, such that the at-risk groups showed slower initial growth followed by accelerated growth during the last two thirds of the year (see Figure 5). Most important was that the two intervention groups demonstrated quicker growth rates than the enhanced classroom group and the typically achieving group as evidenced by steeper slopes (see last column of Table 3).

Effect Sizes. We also examined the effects of our interventions in terms of effect sizes on the intercept and slope terms centered at the final wave of growth assessments. This helped us quantify the magnitude of differences on the intercept (i.e., April scores) and slope (i.e., growth rates) of the two supplemental intervention groups relative to the intercept and slope of the enhanced classroom group. Table



4 shows that effect sizes for the intercept, or April scores, tended to be large for both interventions and for untimed word reading, word reading fluency, and phonological awareness. Effect sizes on phonological awareness and untimed word reading were somewhat larger for the Proactive than for the Responsive intervention, while effect sizes for April scores in word reading fluency were slightly larger for the Responsive intervention. For April scores in passage reading fluency (cohort 1) effects were small for the Proactive group and large for the Responsive Group. For the second cohort, effect sizes were large for both groups, but they were somewhat larger for the Responsive group. The effects of Proactive intervention on April scores in nonword reading fluency were moderate, whereas the effects of Responsive intervention on this outcome were small.

The effect sizes for the Proactive intervention for growth (slope) in phonological awareness and word reading fluency were large, while effects on growth in untimed word reading were moderate and effects on growth in nonword reading fluency and passage reading fluency were small. The Responsive intervention demonstrated moderate effects on growth in word reading fluency and smaller effects on growth on all other measures. The overall pattern of effect sizes is consistent with the different instructional emphases of the two programs.

Analyses of End-of-Year Achievement

Multilevel modeling was used to control for classroom effects and examine intervention-group differences on end-of-year (May) literacy scores. WJ-III achievement was analyzed using *W*-scores rather than age-based or grade-based standard scores because we were interested in examining group differences in absolute abilities rather than relative abilities as a norm-referenced score would produce. Examination of potential covariates found significant classroom effects on WJ-III Spelling, WJ-III Calculations, and CRAB-R Fluency, $\chi^2(2) = 1.71, 2.33, 1.81, ps < .05$; significant main effects of ethnicity on WJ-III Word Attack, $F(2, 303) = 5.80, p < .01$; and significant main

effects of gender on CRAB-R Fluency, $F(1, 296) = 5.96, p < .05$. After controlling for classroom, ethnicity, and gender when appropriate, we found significant overall group differences on all end-of-year literacy scores, $F(3, 262-305) = 13.45$ to $23.26, ps < .001$, and end-of-year WJ-III Calculation scores, $F(3, 304) = 9.15, p < .001$.

Follow-up contrasts examined the effectiveness of the interventions in terms of effects on end-of-year achievement scores (see Table 5). These contrasts demonstrated that the two intervention groups obtained significantly higher scores than the enhanced classroom group on end-of-year WJ-III Word Identification and Spelling. Only the Proactive group demonstrated significantly higher differences than the enhanced classroom group on WJ-III Word Attack. Neither of the interventions demonstrated significant effects on WJ-III Calculations, WJ-III Reading Fluency, WJ-III Passage Comprehension, CRAB-R Fluency, or CRAB-R Comprehension. Table 6 presents *W*-scores and grade-based standard scores on the WJ-III outcomes or raw scores on the CRAB-R outcomes for each group, as well as relevant effect sizes.

Analyses of the WJ-III Reading Fluency and both CRAB-R subtests included only children who read well enough to be administered these tests. There were 43 children who were administered the end-of-year assessment battery but who did not read well enough to be administered the WJ-III Reading Fluency test. Moreover, the four groups differed in terms of the proportion of children who could or could not be administered WJ-III Reading Fluency at the end of the year, $\chi^2(3) = 19.14, p < .001$. Specifically, 2 of the 43 children were in the typically achieving group, 7 children were in the Proactive intervention group, 14 children were in the Responsive intervention group, and 20 children were in the enhanced classroom group. The difference in proportions of children who were testable between the Proactive and enhanced classroom groups was significant, $\chi^2(1) = 5.29, p = .02$. Similarly, there were eight children who were administered the

Table 5
Group Differences in End of the Year Achievement

Contrasts	WJ-III Word Attack W-Score				WJ-III Word Identification W-Score				WJ-III Passage Comprehension W-Score				WJ-III Reading Fluency W-Score			
	diff.	(SE)	df	t	diff.	(SE)	df	t	diff.	(SE)	df	t	diff.	(SE)	df	t
P vs. EC	11.00	(2.74)	303	4.01***	10.81	(3.26)	305	3.31***	3.19	(2.35)	305	1.36	0.01	(2.57)	262	0.00
R vs. EC	4.11	(2.70)	303	1.53	7.58	(3.21)	305	2.36*	4.50	(2.31)	305	1.95+	3.35	(2.59)	262	1.29
P vs. R	6.88	(2.80)	303	2.45*	3.23	(3.33)	305	0.97	-1.31	(2.40)	305	-0.54	-3.34	(2.59)	262	-1.29
P vs. TA	4.35	(2.76)	303	1.59	-13.92	(3.24)	305	-4.30***	-12.77	(2.33)	305	-5.47***	-14.03	(2.42)	262	-5.79***
R vs. TA	-11.20	(2.67)	303	-4.20***	-17.15	(3.18)	305	-5.39***	-11.46	(2.29)	305	-5.00***	-10.68	(2.44)	262	-4.38***
Contrasts	WJ-III Spelling W-Score				WJ-III Calculations W-Score				CRAB-R Fluency W-Score				CRAB-R Comprehension W-Score			
	diff.	(SE)	df	t	diff.	(SE)	df	t	diff.	(SE)	df	t	diff.	(SE)	df	t
P vs. EC	8.13	(2.37)	305	3.42***	1.00	(1.83)	304	0.55	6.49	(3.92)	269	1.66+	0.45	(.57)	304	0.79
R vs. EC	8.22	(2.33)	305	3.45***	2.00	(1.80)	304	1.11	6.83	(3.88)	269	1.76+	0.91	(.56)	304	1.62
P vs. R	-0.09	(2.42)	305	-0.04	1.00	(1.87)	304	0.54	-0.34	(3.97)	269	-0.09	-0.46	(.58)	304	-0.79
P vs. TA	-8.37	(2.36)	305	-3.54***	7.38	(1.82)	304	4.05***	-19.02	(3.86)	269	-4.93***	-3.49	(.57)	304	-6.15***
R vs. TA	-8.28	(2.31)	305	-3.58***	8.38	(1.78)	304	4.07***	-18.67	(3.79)	269	-4.93***	-3.02	(.56)	304	-5.46***

Note: P = Proactive; R = Responsive; EC = enhanced classroom; TA = typically achieving; diff = difference between the estimates; df = degrees of freedom. +p < .10; * p < .05; **p < .01; ***p < .001.

Table 6
End-of-Year Outcome Measure by Group

Measure	Proactive		Responsive		Enhanced Classroom		Typically Achieving		ES Proactive vs. EC	ES Responsive vs. EC				
	W	(SE)	M	W	(SE)	M	W	(SE)			M			
WJ-III ^a														
Word Attack	480.80	(2.16)	109	473.92	(2.10)	105	469.81	(2.03)	103	485.12	(1.99)	111	0.76	0.28
Word Identification	446.75	(2.43)	107	443.50	(2.36)	106	435.94	(2.26)	102	460.67	(2.22)	113	0.51	0.36
Passage Comprehension	456.80	(1.82)	95	485.11	(1.76)	97	453.61	(1.70)	94	469.57	(1.66)	103	0.21	0.30
Spelling	461.64	(1.94)	102	461.73	(1.88)	103	453.51	(1.83)	97	470.01	(1.78)	109	0.54	0.55
Calculations	463.02	(1.61)	104	462.02	(1.57)	103	464.02	(1.53)	105	470.40	(1.49)	111	0.09	0.17
Reading Fluency ^b	441.85	(1.91)	98	445.20	(1.93)	100	441.85	(1.90)	98	445.88	(1.69)	107		
CRAB-R ^c														
Fluency		(3.60)	44		(3.60)	44		(3.50)	37		(3.40)	63		
Comprehension		(0.47)	4.3		(0.46)	4.7		(0.45)	3.8		(7.70)	7.7		

Note. M = WJ-III age-based standard score or CRAB-R raw scores.

CRAB-R and WJ-III Reading Fluency effect sizes are not reported because of the inability of some students to take these subtests resulting in biasing outcomes in favor of the enhanced classroom condition.

Sample Size:

^a Proactive *n* = 80 Responsive *n* = 83 Enhanced classroom *n* = 82 Typically achieving *n* = 98
^b Proactive *n* = 73 Responsive *n* = 69 Enhanced classroom *n* = 62 Typically achieving *n* = 96
^c Proactive *n* = 80 Responsive *n* = 80 Enhanced classroom *n* = 77 Typically achieving *n* = 98

end-of-year assessment battery but who could not be administered either of the CRAB-R subtests. Of these children, three participated in the Responsive intervention, and five participated in the enhanced classroom condition.

The two intervention groups had comparable scores on all but one end-of-year outcome measure. Specifically, the Proactive group had higher scores than the Responsive group on end-of-year WJ-III Word Attack skills. The Proactive intervention effect on word attack was also the only exception to the general finding that children in the intervention groups continued to have lower end-of-year scores than children in the typically achieving group.

Effect sizes for the intervention groups relative to the enhanced classroom group were calculated in the same manner as for the growth analyses, except that effect sizes were only calculated for the intercept because there was no slope term. It should also be noted that the scores on the norm-referenced tasks are largely in the average range. Effect sizes are in the small to moderate range (see Table 6). Effects sizes are largely comparable in the two intervention groups except for word attack, where effects for Proactive were larger.

Hypothesis 4: Child Characteristics Predicting Response to Intervention

We examined hypothesis 4 by asking if children's initial status on relevant reading-related skills interacted with group when predicting either growth rates or April scores of growth outcomes or end-of-year standardized achievement. The interaction terms that specifically addressed hypothesis 4 were sequentially dropped if found nonsignificant so that any main effects of initial reading skills on literacy acquisition would be apparent.

Growth Outcomes

Phonological Awareness. Students' phonological awareness at the beginning of the year was significantly and positively related to all

fluency scores across the school year (see main effects of phonological awareness on intercepts for TOWRE word reading fluency, TOWRE nonword reading fluency, and TCARE passage reading fluency in Table 7). Initial phonological awareness scores were positively associated with growth in all fluency scores, such that higher initial phonological awareness scores were associated with steeper growth in fluency and lower initial phonological awareness scores were associated with slower growth in fluency (see main effects of phonological awareness on slopes in Table 7). The positive relations of initial phonological awareness with growth in fluency and with fluency scores throughout the year were most pronounced in the typically achieving group and were less pronounced in the three at-risk groups, which accounted for the significant group by phonological awareness interaction effects in Table 7. No significant interactions were apparent among the at-risk groups.

The relationship of initial phonological awareness with growth in untimed word reading was very different than the relationship of initial phonological awareness with growth in fluency. Initial phonological awareness was still positively related with untimed word reading scores at all time points; however, initial phonological awareness was *negatively* related to growth in untimed word reading in the typically achieving group. The negative relationship of initial phonological awareness with growth among the typically achieving students, along with the relative lack of association between initial phonological awareness and growth in the at-risk groups, accounted for the significant group by phonological awareness interaction on the slope term of the untimed word reading model in Table 7.

Efficiency of Phonological Access. Letter naming efficiency on RAN at the beginning of the year was significantly and positively related to fluency scores across the school year (see main effects of RAN on intercepts for word reading fluency, nonword reading fluency, and passage reading fluency in Table 8).

Table 7
Effects of Initial Phonological Awareness on Growth in Reading and Response to Intervention

	Intercept		Slope		Quadratic	
	df	F	df	F	df	F
Word reading fluency						
Group	3, 85	19.05***	3, 1423	5.47***	3, 1423	3.45*
PA	1, 1423	46.35***	1, 1423	6.58***	1, 1419	1.34 ^a
Group X PA	3, 1423	6.15***	3, 1420	2.34 ^a	3, 1416	1.19 ^a
Nonword reading fluency						
Group	3, 85	7.56***	3, 1431	4.06**	—	—
PA	1, 1431	62.32***	1, 1431	10.11**	—	—
Group X PA	3, 1431	13.61***	3, 1428	0.65 ^a	—	—
Passage reading fluency (cohort 1)						
Group	3, 56	4.55**	3, 1924	2.68*	3, 1924	1.58 ^a
PA	1, 1924	15.21***	1, 1924	9.12**	1, 1923	0.37 ^a
Group X PA	3, 1924	0.89	3, 1924	2.66*	3, 1920	0.58 ^a
Passage reading fluency (cohort 2)						
Group	3, 53	7.32***	3, 2365	6.62***	3, 2365	13.92***
PA	1, 2365	28.33***	1, 2365	10.47***	1, 2361	0.04 ^a
Group X PA	3, 2365	10.88***	3, 2362	0.28 ^a	3, 2358	1.14 ^a
Untimed word reading						
Group	3, 85	5.07**	3, 1420	37.17***	—	—
PA	1, 1420	44.33***	1, 1420	19.53***	—	—
Group X PA	3, 1420	2.51	3, 1420	10.69***	—	—

Note. *df* = degrees of freedom; PA = phonological awareness. Models of growth in untimed word reading and nonword reading fluency did not include quadratic terms.

^aNonsignificant terms that were dropped from the model, resulting in an increase in degrees of freedom.

* $p < .05$; ** $p < .01$; *** $p < .001$.

In addition, initial RAN scores were positively associated with growth in all fluency scores. Higher initial RAN scores were associated with steeper growth in fluency, and lower initial RAN scores were associated with slower growth in fluency (see main effects of RAN on fluency slopes in Table 8). Moreover, initial RAN scores significantly interacted with group when predicting growth patterns in fluency (see Table 8). Specifically, initial RAN scores were more strongly related to fluency scores at all testing periods in the typically achieving group than in the three at-risk groups, where no interactions were apparent.

The relationship of initial RAN with untimed word reading was similar to that of initial phonological awareness with word reading. Initial RAN was positively related to untimed word reading scores at all time points; however, initial RAN was *negatively* related to growth rates in the typically achieving group, accounting for the significant group by RAN interaction on the slope term of the untimed word reading model in Table 8. Again, there were no interactions apparent among the at-risk groups.

Vocabulary. Table 9 shows that WASI Vocabulary scores were unrelated to growth

Table 8
Effects of Initial RAN on Growth in Reading and Response to Intervention

	Intercept		Slope		Quadratic	
	df	F	df	F	df	F
Word reading fluency						
Group	3, 85	3.57*	3, 1419	6.12***	3, 1419	3.39*
RAN	1, 1419	89.65***	1, 1419	18.91***	1, 1415	0.25 ^a
Group X RAN	3, 1419	10.96***	3, 1416	2.59 ^a	3, 1412	0.33 ^a
Nonword reading fluency						
Group	3, 85	8.76***	3, 1427	5.04**	—	—
RAN	1, 1427	58.01***	1, 1427	19.36***	—	—
Group X RAN	3, 1427	16.25***	3, 1424	0.20 ^a	—	—
Passage reading fluency (cohort 1)						
Group	3, 56	0.48	3, 1930	1.82	3, 1927	1.51 ^a
RAN	1, 1930	30.17***	1, 1930	18.24***	1, 1923	0.18 ^a
Group X RAN	3, 1930	2.65*	3, 1924	0.04 ^a	3, 1920	0.11 ^a
Passage reading fluency (cohort 2)						
Group	3, 53	0.82	3, 2349	1.44	3, 2349	14.04***
RAN	1, 2349	42.39***	1, 2349	14.62***	1, 2348	2.74 ^a
Group X RAN	3, 2349	1.47	3, 2349	3.73**	3, 2345	0. ^a
Untimed word reading						
Group	3, 85	4.14**	3, 1416	2.03	—	—
RAN	1, 1416	32.99***	1, 1416	17.00***	—	—
Group X RAN	3, 1416	1.41	3, 1416	4.40**	—	—

Note. *df* = degrees of freedom; RAN = rapid letter naming. Models of growth in untimed word reading and nonword reading fluency did not include quadratic terms.

^aNonsignificant terms that were dropped from the model, resulting in an increase in degrees of freedom.

p* < .05; *p* < .01; ****p* < .001.

patterns in word reading fluency, nonword reading fluency, and passage reading fluency in the second cohort. However, WASI Vocabulary was positively related to growth rates on passage reading fluency among children in the Proactive and typically achieving groups who were in the first cohort. In other words, children in these two groups with higher vocabularies improved their passage reading fluency faster than children in the same groups with smaller vocabularies. WASI Vocabulary scores were unrelated to growth rates on the Passage Reading Fluency subtest among children in the Responsive or

enhanced classroom groups. The significant group by vocabulary interaction described above indicated that children's vocabularies played a greater role in passage reading fluency growth if children were in the typically achieving or Proactive groups than if children were in the Responsive or enhanced classroom groups. In terms of vocabulary effects on growth in untimed word reading, there were only main effects of vocabulary such that children with higher vocabulary scores generally scored higher on untimed word reading across the year regardless of which group they were in.

Table 9
Effects of Initial WASI Scores on Growth in Reading and Response to Intervention

	Intercept		Slope		Quadratic	
	df	F	df	F	df	F
Word reading fluency						
Group	3, 85	36.08***	3, 1428	4.53**	3, 1428	3.46*
WASI	1, 1284	0.76 ^a	1, 1280	0.00 ^a	1, 1276	0.38 ^a
Group X WASI	3, 1281	0.84 ^a	3, 1277	0.04 ^a	3, 1273	0.41 ^a
Nonword reading fluency						
Group	3, 85	18.11***	3, 1436	2.14	—	—
WASI	1, 1292	2.08 ^a	1, 1288	0.68 ^a	—	—
Group X WASI	3, 1289	0.58 ^a	3, 1285	0.24 ^a	—	—
Passage reading fluency (cohort 1)						
Group	3, 55	0.87	3, 1862	1.66	3, 1862	3.07*
WASI	1, 1862	1.68	1, 1862	1.76	1, 1862	1.25
Group X WASI	3, 1862	0.76	3, 1862	2.25	3, 1862	3.32*
Passage reading fluency (cohort 2)						
Group	3, 53	16.45***	3, 2370	4.86**	3, 2370	13.98***
WASI	1, 2015	0.17 ^a	1, 2011	0.36 ^a	1, 2007	2.31 ^a
Group X WASI	3, 2012	1.80 ^a	3, 2008	2.03 ^a	3, 2004	2.28 ^a
Untimed word reading						
Group	3, 79	11.37***	3, 1286	44.16***	—	—
WASI	1, 1286	6.49**	1, 1285	1.53 ^a	—	—
Group X WASI	3, 1282	0.03 ^a	3, 1279	2.32 ^a	—	—

Note. *df* = degrees of freedom; WASI = Vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence. Models of growth in untimed word reading and nonword reading fluency did not include quadratic terms.

^aNonsignificant terms that were dropped from the model, resulting in an increase in degrees of freedom.

p* < .05; *p* < .01; ****p* < .001.

End-of-Year Outcomes

In terms of norm-referenced end-of-year outcomes, analyses of group by child characteristic interactions found that initial phonological awareness was differentially important for children in the various intervention groups when predicting end-of-year word attack skills (see Table 10). Specifically, initial phonological awareness scores were less closely related to end-of-year word attack scores in the Proactive group relative to all other groups. For all other end-of-year outcomes, initial phonological awareness scores only had significant main effects, such that children with higher

phonological awareness scores in October had higher literacy scores in May. Initial RAN scores only demonstrated significant main effects (see Table 11), such that higher RAN scores in October were associated with higher literacy scores in May. In a similar manner, there were only main effects of vocabulary scores (see Table 12), such that higher vocabulary scores were associated with higher end-of-year literacy scores.

Discussion

The purpose of this study was to address the effectiveness of combining enhanced

Table 10
Effects of Initial Phonological Awareness on End-of-Year Outcomes and Responses to Intervention

	WJ-III Word Attack		WJ-III Word Identification		WJ-III Passage Comprehension		WJ-III Reading Fluency	
	df	F	df	F	df	F	df	F
Group	3, 299	2.30	3, 304	10.22***	3, 304	6.87***	3, 261	5.16***
PA	1, 299	50.70***	1, 304	56.70***	1, 304	42.80***	1, 261	31.72***
Group X PA	3, 299	2.72*	3, 301	0.86 ^a	3, 301	0.79 ^a	3, 258	1.02 ^a
	WJ-III Spelling		WJ-III Calculations		CRAB-R Fluency		CRAB-R Comprehension	
	df	F	df	F	df	F	df	F
Group	3, 302	10.62***	3, 303	1.22	3, 295	8.55***	3, 303	7.65***
PA	1, 304	23.72***	1, 303	49.09***	1, 295	15.85***	1, 303	42.29***
Group X PA	3, 301	0.31 ^a	3, 300	1.24 ^a	3, 292	0.42 ^a	3, 300	1.79 ^a

Note. WJ-III = Woodcock–Johnson-III; CRAB-R = Comprehensive Assessment of Reading Battery—Revised; PA = phonological awareness; *df* = degrees of freedom.

^aNonsignificant terms that were dropped from the model, resulting in an increase in degrees of freedom.

p* < .05; *p* < .01; ****p* < .001.

Table 11
Effects of Initial RAN on End-of-Year Outcomes and Responses to Intervention

	WJ-III Word Attack		WJ-III Word Identification		WJ-III Passage Comprehension		WJ-III Reading Fluency	
	df	F	df	F	df	F	df	F
Group	3, 301	7.52***	3, 303	8.85***	3, 304	6.91***	3, 261	5.00**
RAN	1, 301	14.97***	1, 303	38.02***	1, 304	25.52***	1, 261	30.11***
Group X RAN	3, 298	0.93 ^a	3, 300	2.44 ^a	3, 301	1.89 ^a	3, 258	0.83 ^a
	WJ-III Spelling		WJ-III Calculations		CRAB-R Fluency		CRAB-R Comprehension	
	df	F	df	F	df	F	df	F
Group	3, 303	7.94***	3, 302	2.34	3, 294	4.33**	3, 302	5.84***
RAN	1, 303	38.25***	1, 302	19.26***	1, 294	58.65***	1, 302	48.34***
Group X RAN	3, 300	1.40 ^a	3, 299	1.98 ^a	3, 291	2.45 ^a	3, 299	0.65 ^a

Note. WJ-III = Woodcock–Johnson-III; CRAB-R = Comprehensive Assessment of Reading Battery—Revised; RAN = rapid automatic naming of letter names; *df* = degrees of freedom.

^aNonsignificant terms that were dropped from the model, resulting in an increase in degrees of freedom.

p* < .05; *p* < .01; ****p* < .001.

classroom reading instruction with small-group supplemental reading instruction derived from either behavioral or cognitive theory for first-grade students at risk for reading difficulties. Specifically, we hypothesized (a) that

small-group reading instruction, in the form of the Responsive and Proactive interventions, provided in addition to the classroom reading program, would be more effective than high-quality classroom reading instruction alone

Table 12
Effects of Initial Vocabulary on End-of-Year Outcomes and Responses to Intervention

	WJ-III Word Attack		WJ-III Word Identification		WJ-III Passage Comprehension		WJ-III Reading Fluency	
	df	F	df	F	df	F	df	F
Group	3, 274	9.04***	3, 276	14.98***	3, 276	12.57***	3, 237	9.35***
WASI	1, 274	14.82***	1, 276	6.94**	1, 276	12.11***	1, 237	16.30***
Group X WASI	3, 271	0.96 ^a	3, 273	0.31 ^a	3, 273	0.30 ^a	3, 234	0.18 ^a

	WJ-III Spelling		WJ-III Calculations		CRAB-R Fluency		CRAB-R Comprehension	
	df	F	df	F	df	F	df	F
Group	3, 276	14.12***	3, 275	6.65***	3, 295	15.03***	3, 303	19.11***
WASI	1, 276	4.12*	1, 275	10.13***	1, 295	116.20***	1, 303	94.65***
Group X WASI	3, 273	0.69 ^a	3, 272	0.75 ^a	3, 292	0.56 ^a	3, 300	0.55 ^a

Note. WJ-III = Woodcock-Johnson-III; WASI = WASI vocabulary; CRAB-R = Comprehensive Assessment of Reading Battery—Revised; *df* = degrees of freedom.

^aNonsignificant terms that were dropped from the model, resulting in an increase in degrees of freedom.

p* < .05; *p* < .01; ****p* < .001.

for students at risk for reading failure; (b) that these two interventions would be comparably effective; (c) that the reading performance of at-risk students who received this additional intervention would approach the level of performance of their normally developing peers; and (d) that specific child characteristics would differentially predict individual responses to these two interventions, which were provided in different formats and emphasized different aspects of the reading process.

Our results revealed that struggling first-grade readers who received one of the two interventions did, on average, perform better on multiple measures of reading after participating in either Responsive or Proactive interventions than children who received only enhanced classroom instruction (hypothesis 1). These gains were true both in terms of rate of growth (slope) and in terms of end-of-year status (intercept). Specifically, the students who received either form of supplemental intervention were better able to (a) read words in both timed and untimed formats, (b) spell words, (c) read connected text accurately and fluently, and (d) demonstrate phonological awareness than were the at-risk readers who did not receive

intervention. Further, consistent with hypothesis 2, the outcomes achieved by each intervention were largely comparable as evidenced by similar overall average effect sizes (Proactive *ES* = .84 and Responsive *ES* = .78). In terms of hypothesis 3, these outcomes reflect the more rapid growth of Proactive and Responsive students relative to the enhanced classroom group and, in some cases, the typically achieving group. Even so, struggling readers who participated in our interventions did not attain academic levels commensurate with their normally achieving peers in the same classrooms receiving the same core reading instruction, with the one exception of word attack ability of children in the Proactive condition. However, the majority of children at risk for reading failure who participated in intervention did reach average achievement levels on many normative measures. In terms of hypothesis 4, there was little evidence for interactions between child characteristics and intervention programs.

Hypothesis 1: Added Value of Small-Group Supplemental Intervention

To fully appreciate the impact of the interventions, it is important to first consider the

effectiveness of the enhanced classroom condition, which promoted high levels of reading growth for many children at risk for reading failure. Although we cannot separate the efficacy attributable to the districts' professional program and our assessment and consultation additions to classroom programs, only 16% of at-risk readers in our sample who received enhanced classroom instruction alone remained below average performance levels on basic reading skills at the end of first grade. Extrapolating to the total school population, this figure translates to only 3% of all children. In other studies of classroom-level instruction, inadequate responder rates have only been reduced to 5% to 7% (e.g., Denton & Mathes, 2003; Mathes & Denton, 2002). Likewise, at-risk readers in the enhanced classroom group achieved, on average, standard scores that placed them consistently in the average range on multiple measures by the end of first grade. Thus, it is fair to say that the enhanced classroom condition served as a rigorous comparison group for the two interventions.

Even so, regardless of the nature of the small-group intervention, children who received supplemental small-group intervention performed significantly better than their at-risk peers who received only enhanced classroom instruction on tests of phonological awareness, timed and untimed word reading, passage reading fluency, and spelling. These findings are educationally significant, given the close association of phonological awareness (Wagner, 1988) and accurate and fluent word reading (Lyon, 1995) with successful reading acquisition and development in later grades (Torgesen & Burgess, 1998) and the close association of fluent text reading with reading comprehension (Fuchs, Fuchs, Hosp, & Jenkins, 2001).

Further, students in both Proactive and Responsive groups had significantly faster rates of learning than those of the typically achieving comparison group in phonological awareness, word reading, and passage reading fluency of first-grade text (cohort 2). These results indicate that both interventions promoted a "closing of the gap" for students who began the year

significantly behind their normally developing peers. In both interventions, not only were students developing key skills in phonological awareness and word identification more quickly than children without reading difficulties, but also this development of key reading subskills was manifested in a concurrent growth in the ability to read end-of-year connected text smoothly and accurately. In fact, the average end point of the students in both interventions of approximately 55 words read correctly per minute places them firmly within average parameters for passage reading fluency (Good, Wallin, Simmons, Kame'enui, & Kaminski, 2002). Beyond the issue of statistical significance, the magnitude of the differences between intervention groups and the enhanced classroom condition on most measures was moderate to large, and thus educationally relevant and meaningful. These effect sizes are relative to the gains of the enhanced classroom group, representing the magnitude of the *value-added* impact of the interventions, rather than effect sizes comparing intervention to a no-intervention control.

Perhaps the most compelling evidence of the value-added impact of the two reading interventions comes from simple counts of the number of children reading below the average range. We adopted a cut point used in many early intervention studies (Torgesen, 2000; Torgesen et al., 2003), namely performance below the 30th percentile on the Woodcock-Johnson III Basic Reading Skills cluster. The 30th percentile is an arbitrary designation, representing the lower end of the average range. If we define *average* conventionally as all children with one standard deviation of the mean, the range on the WJ-III would be 92.5–107.5. Developing a confidence interval that takes into account the standard error of measurement (4 points) would yield a range of 92 (30th percentile) to 108. Obviously we can compute these outcomes using other measures, but the results are similar across domains. Using the Basic Reading cut point, 16% (15 of 92 students) in the enhanced classroom instruction condition, 7% (6 of 83 students) in the Responsive intervention, and 1% (1 of 80 students) in the

Proactive group were reading below the average range at the end of the intervention period. Extrapolating these figures to the total population, based on the fact that the TPRI identifies the lowest 20% of readers, enhanced classroom instruction reduced the number of poor readers to about 3%, Responsive and Proactive to below 1%. Thus, all three instructional conditions significantly reduced the number of at-risk children, but they had increased benefit from the supplemental interventions. Given that 6% of students across the nation receive special education services for a learning disability (over 90% with reading problems; Kavale & Reese, 1992), and that many more students struggle in becoming readers, these reductions are educationally significant. Students in the at-risk groups were predominantly minority, economically disadvantaged students, and hardly restricted to those at risk for special education placement.

Hypothesis 2: Comparison of the Two Interventions

Although both interventions were associated with better outcomes than enhanced classroom instruction, there was little reliable evidence that the two interventions were differentially effective. Overall average effect sizes also illustrate that these interventions were comparable in their impact.

At the same time, intervention specificity was apparent when examining outcomes in terms of both statistical significance and effect size. For example, students who received Proactive intervention improved their word reading fluency and nonword reading fluency more quickly than students who received only classroom instruction, whereas students who received Responsive intervention improved their word reading fluency and nonword reading fluency at the same rate as students who received only enhanced classroom instruction. In a similar manner, students in the Proactive intervention also had better word attack skills at the end of the year than students who received only enhanced classroom instruction. Students

in the Responsive intervention had end-of-year word attack skills that were comparable to those of students who received only enhanced classroom instruction. In contrast, effect sizes for oral reading fluency were higher for the Responsive intervention than for Proactive, congruent with the relatively higher percentage of time Responsive reading students spent engaged in reading connected text. These findings lend credibility to the idea that students learn what they are taught (Allington, 1983), and they support the validity of the differences in the two interventions, as the Proactive intervention placed greater emphasis on phonological awareness, sounding out words in isolation, and reading words in lists, and students in the Responsive intervention spent relatively more of their lesson time reading connected text. Even so, to put these differences into perspective, the two interventions yielded comparable scores on seven of eight end-of-year outcomes. Furthermore, the average performance of students in both interventions was in the average range in terms of standard scores.

Hypothesis 3: Closing the Achievement Gap

Students who participated in either Proactive or Responsive interventions had steeper slopes than typically achieving students on several measures, including untimed word reading, phonological awareness, and passage reading fluency (cohort 2). Further, students who received the Proactive intervention had steeper slopes on word reading fluency and nonword reading fluency than typically achieving students. These findings suggest that intervention students were learning more rapidly than their typically developing peers, thereby progressively closing the achievement gap. However, there was only one instance in which the achievement gap was fully closed. This was achieved in May by the Proactive group on word attack skills. Of course, it should also be recognized that the typically achieving students in this sample scored in the high average range on all but one posttest measure (see Table 6).

Thus, despite scores in the average or better range, it is not surprising that our at-risk students did not achieve statistical equivalence with their higher performing peers given their high performance, another testament to the quality of instruction in the six schools.

Hypothesis 4: The Interaction of Child Characteristics With Responsiveness to Intervention

Contrary to our predictions, we found that, in general, at-risk readers who started the year with low levels of phonological awareness, less efficient letter-naming ability, or lower vocabulary knowledge were equally likely to make progress regardless of their assignment to the Proactive or Responsive groups. The only exceptions to our findings were that initial phonological awareness had less impact on WJ-III Word Attack scores for students who received the Proactive intervention and that initial vocabulary scores were less important to Passage Reading Fluency outcomes for cohort 1. While consistent with our hypotheses, it should be noted that these findings were not replicated on other measures testing similar constructs and, in the case of the Passage Reading Fluency subtest, were not replicated with cohort 2. Further, the interactions were not wide reaching, with no interaction being detected for any other measures regardless of initial status scores. Thus, it is fair to say that, while initial status on phonological awareness, phonological access, and vocabulary predicted outcomes, interactions with the efficacy of the interventions were not apparent.

This finding has important implications for educators who must select from multiple research-supported early reading interventions. It appears that different interventions can be effective for a wide variety of students as long as they include key components addressing phonological awareness, phonological decoding, fluency in reading words in isolation and in text, and comprehension of text. Further, the way in which the key components are taught to students can differ. It is important to remember

that the teachers in this research were selected because they possessed personal philosophies aligning them to either Proactive or Responsive methods. Having this initial “buy in” ensured that these teachers believed in the approach to which they were assigned and worked hard to deliver daily instruction to the best of their ability. Had we taken a one-method approach, we speculate that some of our teachers would have felt alienated, and less willing to deliver daily instruction in accordance with our research protocol. In short, the results of this research suggest that, at the very least, there is room for choice in selecting an approach for providing supplemental intervention.

Nonsignificant Findings

Outcomes for which statistically significant differences were not detected between interventions groups and the enhanced classroom group included TOWRE Nonword Reading Efficiency, WJ-III Reading Fluency, WJ-III Passage Comprehension, WJ-III Calculations, CRAB-R Fluency, and CRAB-R Comprehension. We suspect that the lack of significant differences between at-risk groups on the TOWRE Nonword Reading Efficiency subtest is related to speed of phonological access. Because the TOWRE is a timed test, students had to quickly access their knowledge of letter–sound correspondences and quickly string the sounds together to form words, a task requiring both phonological access and adequate working memory. The lack of statistical differences likely reflects problems with phonological access because on a similar but untimed measure (i.e., WJ-III Word Attack) a significant difference was detected for the Proactive intervention group. Alphabetic knowledge and working memory demands were identical on these two tasks; however, speeded phonological access demands were substantially higher on the TOWRE.

The WJ-III subtests for which the year-end analysis indicated a lack of statistically significant differences favoring the intervention groups were reading fluency, passage comprehension, and mathematical calculations. When

examining the results of the WJ-III Reading Fluency subtest and the CRAB-R Fluency and Comprehension subtests, it is important to keep in mind that children who could not complete the sample items were excluded from this subtest, resulting in removal of the lowest performers and increasing mean scores. Because more children in the enhanced classroom group were excluded from this subtest, mean scores were biased in favor of the enhanced classroom group, making them appear to be more similar to the intervention groups. The fact that significantly more students in the Proactive intervention were actually able to be administered this subtest demonstrated the advantage of this intervention. It is important to keep in mind that growth in passage reading fluency on first-grade text, which had a lower floor, did yield significant differences favoring the two intervention groups (cohort 2), thus corroborating the advantage of the interventions on fluency development.

In terms of calculations, we did not expect to find differences because we did not intervene in the domain of mathematics. If students in the intervention groups were exhibiting generalized growth rather than development specific to reading acquisition, they might be expected to differ from other at-risk students in their ability to perform mathematical calculations. This was not the case, indicating that growth was specific to the domain in which they received intervention.

Implications for Practice

This research affirms the value of providing early reading intervention to struggling readers. Students who participated in one of the two interventions, on average, finished first grade better prepared for second grade than students who received only enhanced classroom instruction. Further, this research lends support to our hypothesis that intervention instruction should be provided in tandem with quality classroom instruction. In other studies with similar types of interventions that were not provided in conjunction with quality classroom instruction, the results, while positive, have not reduced the

levels of struggling readers to levels demonstrated in the current research. Likewise, this research clearly demonstrates that enhanced classroom instruction alone is inadequate for a small number of students who require instruction of higher intensity.

Perhaps the most important finding of this research is that supplemental intervention approaches derived from different theoretical perspectives were both effective. These findings suggest to us that there is likely not “one best approach” and not one right philosophy or theory for how to best meet the needs of struggling readers. Nor did we find evidence that one approach was better for some at-risk children than another. Thus, the outcomes of this research led us to surmise that schools and teachers can be granted some latitude in choosing an approach to providing supplemental instruction for struggling readers. We hypothesize that if schools are allowed to choose from among effective choices an approach that best aligns to personal philosophy and theory, then there is likely to be less resistance, higher quality implementation, as well as sustainability over the long term.

At the same time it is critical that our outcomes *not* be interpreted as saying that the content included in supplemental instruction for struggling readers does not matter. It is clearly not the case that “anything goes.” Both the Proactive and Responsive Reading interventions included elements that have been identified as critical for instruction of students who struggle to acquire the ability to read well (see Foorman & Torgesen, 2001; NICHD, 2000; Rayner et al., 2001; Snow et al., 1998). Both interventions provided for instruction in key reading skills, balanced with opportunities to apply reading and writing skills in connected text, and they both provided students with explicit instruction and practice in skills related to phonemic awareness, decoding, fluent word recognition and text processing, and spelling. Likewise, both approaches provided instruction in comprehension strategies applied to connected text. Thus, the interventions were comprehensive, integrated approaches to reading instruction.

This research also indicates that it is possible to provide effective early reading instruction to students at risk for reading difficulties using text that is not phonetically decodable and without following a detailed scope and sequence. Teachers in the Responsive Reading intervention in our study had extensive training on topics such as the order in which phonic elements should be introduced (i.e., more useful elements introduced earlier, separating potentially confusing elements such as *b* and *d*), the most frequently occurring words that should be taught at sight, and interpreting the results of assessments and anecdotal observations to plan instruction. In addition, these teachers did not create teaching activities and strategies but rather selected activities from a menu provided to them in the Responsive Reading handbook (Denton & Hocker, 2004). Finally, as described above, students in the Responsive Reading intervention were provided with explicit instruction and practice in phonemic decoding and were taught to rely primarily on graphophonemic information rather than text or picture contexts to identify words. Adding such components to other programs developed from a similar theoretical background may well enhance outcomes with children who have poorer development of prereading skills and in a more cost-effective manner because of implementation in small-group rather than individual instruction.

Study Limitations and Future Directions

In interpreting this research it is important to consider limitations that may reduce the generalizability of our findings. The most significant question involves the provision of 40 minutes of additional reading instruction to students in the supplemental instructional groups. As the first step in our research was to determine whether supplemental instruction had a value-added impact, we did not control for the additional structured reading time. It seems unlikely that simply reading for an additional 40 minutes daily or even extending the language arts block to two hours would produce

similar gains, especially as other studies that have controlled this factor have also shown better gains in small-group instruction relative to classroom instruction alone (Simmons et al., 2003; Torgesen et al., 1999). But the issue of controlling for time spent in reading activities is a next step for this type of research in which supplemental instruction is added to classroom instruction. Second, while conducted in public schools, the interventions were delivered under highly controlled conditions. Even though the intervention teachers were employees of the public schools, they were selected for this research because of their demonstrated expertise as reading teachers. Likewise, we were able to assign teachers to provide the intervention that was most aligned with their personal philosophy and prior teaching experiences. While working with highly knowledgeable and motivated teachers facilitated our ability to test our hypotheses fairly, it is likely that these intervention teachers are not representative of all teachers teaching reading in U.S. schools. Thus, it is not clear if similar results would be achieved with less knowledgeable teachers. Likewise, we conducted this research in relatively high-functioning schools providing solid, core reading instruction. Further research is necessary to determine if similar results would be achieved in schools facing greater challenges.

The authors of the interventions provided considerable coaching and support to the intervention teachers. It is unlikely that similar levels of coaching and support typically would be available to teachers implementing or sustaining these interventions in the “real world.” Thus, it remains to be seen if the results of this research would be replicated under less controlled conditions. Currently, little is known about how best to provide staff development and support to teachers as they work to implement new innovations. Even less is known about maintaining and sustaining innovations (Denton & Fletcher, 2003).

Another area needing further research relates to group size. While group size was held constant at three students per group between the two interventions in this research, future

work needs to be done to determine if other grouping formats, such as one-on-one tutoring, would result in even stronger outcomes, or conversely, if the outcomes can be replicated with larger group sizes. Because group size largely dictates the cost of providing supplemental reading instruction in schools, and because financial considerations can either facilitate or deter implementation of new innovations within schools, the group size issue cannot be ignored.

Conclusion

This study reinforced the added value of supplemental intervention provided to first-grade students who demonstrate risk factors for reading difficulty. Likewise, the results of the study indicate that interventions originating from different theoretical viewpoints, but both of which emphasize word recognition strategies and contain elements previously identified as essential in early reading instruction, can be effective for at-risk first-grade readers. In fact, no reliable interactions were detected between child characteristics and success in one type of intervention or the other.

We propose that these findings lend support to the argument that it is time to stop debating the “best” method for providing early reading intervention. Time is better devoted to determining how to overcome the great challenges that exist in getting effective interventions placed into schools. Likewise, our findings support the idea that schools can be allowed to choose from among good choices those interventions that best fit personal philosophies and personnel talents.

References

- Allington, R.L. (1983). The reading instruction provided readers of differing reading abilities. *Elementary School Journal*, 83, 548–559.
- Becker, W.C. (1973). Applications of behavior principles in typical classrooms. In C.A. Thorenson (Ed.), *Behavior modification in education* (72nd yearbook of the National Society for the Study of Education, part 1, pp. 77–106). Chicago: National Society of the Study of Education.
- Brophy, J., & Good, T. (1986). Teacher-effects results. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York: Macmillan.
- Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18, 32–42.
- Carnine, D.W., Silbert, J., Kame'enui, E.J., & Tarver, S.G. (2004). *Direct instruction reading* (4th ed.). Upper Saddle River, NJ: Merrill-Prentice Hall.
- Clay, M.M. (1987). Learning to be learning disabled. *New Zealand Journal of Educational Studies*, 22, 155–173.
- Clay, M.M. (1993). *Reading Recovery: A guidebook for teachers in training*. Portsmouth, NH: Heinemann.
- Clay, M.M. (2002). *An observation survey of early literacy achievement* (2nd ed.). Portsmouth, NH: Heinemann.
- Denton, C.A., & Fletcher, J.M. (2003). Scaling reading interventions. In B.R. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 445–463). Timonium, MD: York Press.
- Denton, C.A., & Hocker, J. (2004). *Responsive reading instruction: Teacher handbook*. Unpublished manuscript.
- Denton, C.A., & Mathes, P.G. (2003). Intervention for struggling readers: Possibilities and challenges. In B.R. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 229–251). Timonium, MD: York Press.
- Elbaum, B., Vaughn, S., Hughes, M.T., & Moody, S.W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, 92, 605–619.
- Elkonin, D.B. (1973). USSR. In J.A. Downing (Ed.), *Comparative reading: Cross-national studies of behavior and processes in reading and writing* (2nd ed., pp. 551–579). New York: Macmillan.
- Engelmann, S. (1997). *Preventing failure in the primary grades*. Eugene, OR: Association for Direct Instruction.
- Engelmann, S., & Carnine, D. (1982). *Theory of instruction: Principles and applications*. New York: Irvington.
- Foorman, B.R., Francis, D.J., Fletcher, J.M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading disabilities in at-risk children. *Journal of Educational Psychology*, 90, 37–55.
- Foorman, B.R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research and Practice*, 16, 203–212.
- Fountas, I.C., & Pinnell, G.S. (1996). *Guided reading*. Portsmouth, NH: Heinemann.
- Fountas, I.C., & Pinnell, G.S. (1999). *Matching books to readers: Using leveled books in guided reading K–3*. Portsmouth, NH: Heinemann.
- 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, 3–17.
- Fuchs, L.S., Fuchs, D., Hamlett, C.L., & Stecker, P.M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, 28, 617–641.
- Fuchs, L.S., Fuchs, D., Hosp, M.K., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading compe-

- tence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 3, 239–256.
- Gomez-Bellenge, F.X., Rogers, E., & Fullerton, S.K. (2003). *Reading Recovery and Descubriendo la Lectura national report 2001–2002*. Columbus: Ohio State University, Reading Recovery National Data Evaluation Center.
- Good, R.H., Wallin, J., Simmons, D.C., Kame'enui, E.J., & Kaminski, R.A. (2002). *System-wide percentile ranks for DIBELS Benchmark Assessment* (Tech. Rep. 9). Eugene: University of Oregon.
- Harris, K.R., & Pressley, M. (1991). The nature of cognitive strategy instruction: Interactive strategy construction. *Exceptional Children*, 57, 392–404.
- Juel, C. (1988). Learning to read and write: A longitudinal study of children in first and second grade. *Journal of Educational Psychology*, 80, 437–447.
- Kame'enui, E.J., & Simmons, D.C. (1990). *Designing instructional strategies: Prevention of academic learning problems*. Columbus, OH: Merrill.
- Kavale, K.A., & Reese, J.H. (1992). The character of learning disabilities: An Iowa profile. *Learning Disability Quarterly*, 15, 74–94.
- Lyon, R.G. (1995). Toward a definition of dyslexia. *Annals of Dyslexia*, 45, 2–37.
- Mathes, P.G., & Denton, C.A. (2002). The prevention and identification of reading disability. *Seminars in Pediatric Neurology*, 9, 185–191.
- Mathes, P.G., & Fuchs, L.S. (1994). The efficacy of peer tutoring in reading for students with mild disabilities: A best-evidence synthesis. *School Psychology Review*, 23, 55–76.
- Mathes, P.G., Howard, J.K., Allen, S., & Fuchs, D. (1998). Peer-assisted learning strategies for first-grade readers: Making early reading instruction more responsive to the needs of diverse learners. *Reading Research Quarterly*, 33, 62–94. doi:10.1598/RRQ.33.1.4
- Mathes, P.G., & Torgesen, J.K. (1998). All children can learn to read: Critical care for students with special needs. *Peabody Journal of Education*, 73, 317–340.
- Mathes, P.G., Torgesen, J.K., & Allor, J.H. (2001). The effects of peer assisted learning strategies for first-grade learners with and without additional computer assisted instruction in phonological awareness. *American Educational Research Journal*, 38, 371–410.
- Mathes, P.G., Torgesen, J.K., & Heron, J. (2004). Teachers computerized assessment for reading excellence [Computer software]. San Rafael, CA: Talking-Fingers. National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- O'Connor, R. (2000). Increasing the intensity of intervention in kindergarten and first grade. *Learning Disabilities Research & Practice*, 15, 43–54.
- O'Malley, K.J., Francis, D.J., Foorman, B.R., Fletcher, J.M., & SWANK, P.R. (2002). Growth in precursor reading skills: Do low-achieving and IQ-discrepant readers develop differently? *Learning Disability Research and Practice*, 17, 19–34.
- Peterson, B. (1991). Selecting books for beginning readers. In D.E. DeFord, C.A. Lyons, & G.S. Pinnell (Eds.), *Bridges to literacy: Learning from Reading Recovery* (pp. 119–147). Portsmouth, NH: Heinemann.
- Pressley, M.P. (1998). *Elementary reading instruction that works*. New York: Guilford.
- Pressley, M.P., & McCormick, C.M. (1995). *Cognition, teaching, and assessment*. New York: HarperCollins.
- Rayner, K., Foorman, B.R., Perfetti, C.A., Pesetsky, D., & Seidenberg, M.S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, 2, 31–74.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.
- Rojewski, J.W., & Schell, J.W. (1994). Cognitive apprenticeship for learners with special needs. *Remedial and Special Education*, 15, 234–243.
- Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M.C. Wittrock (Ed.), *The handbook of research on teaching* (3rd ed., pp. 376–391). New York: Macmillan.
- Schatschneider, C., Francis, D.J., Foorman, B.R., Fletcher, J.M., & Mehta, P. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology*, 91, 439–449.
- Shanahan, T. (1998). On the effectiveness and limitations of tutoring in reading. *Review of Research in Education*, 23, 217–234.
- Simmons, D.C., Kame'enui, E.J., Stoolmiller, M., Coyne, M.D., & Harn, B. (2003). Accelerating growth and maintaining proficiency: A two-year intervention study of kindergarten and first-grade children at-risk for reading difficulties. In B.R. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 197–228). Baltimore: York Press.
- Skinner, B.F. (1953). *Science and human behavior*. New York: Macmillan.
- Snow, C.E., Burns, M.S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360–397. doi:10.1598/RRQ.21.4.1
- Storch, S.A., & Whitehurst, G.J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, 38, 934–947.
- Torgesen, J.K. (2000). Individual responses in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice*, 15, 55–64.
- Torgesen, J.K., & Burgess, S.R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional studies. In J. Metsala & L. Ehri (Eds.), *Word recognition in beginning reading* (pp. 161–188). Hillsdale, NJ: Erlbaum.
- Torgesen, J.K., Rashotte, C., Alexander, A., Alexander, J., & MacPhee, K. (2003). Progress toward understanding the instructional conditions necessary for remediating reading difficulties in older children. In B.R. Foorman (Ed.), *Preventing and remediating reading difficulties:*

- Bringing science to scale* (pp. 275–298). Baltimore: York Press.
- Torgesen, J.K., Wagner, R.K., Rashotte, C.A., Rose, E., Lindamood, P., Conway, T., et al. (1999). Preventing reading failure in your children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91*, 579–593.
- Vaughn, S., & Linan-Thompson, S. (2003). Group size and time allotted to intervention: Effects for students with reading disabilities. In B.R. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 299–324). Baltimore: York Press.
- Vellutino, F.R., Scanlon, D.M., & Jaccard, J. (2003). Toward distinguishing between cognitive and experiential deficits as primary sources of difficulty in learning to read: A two year follow-up to difficult to remediate and readily remediated poor readers. In B.R. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 73–120). Baltimore: York Press.
- Vellutino, F.R., Scanlon, D.M., Sipay, E.R., Small, S.G., Pratt, A., Chen, R., et al. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88*, 601–638.
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds. & Trans.). Cambridge, MA: Harvard University Press.
- Wagner, R.K. (1988). Causal relations between the development of phonological processing abilities and the acquisition of reading skills: A meta-analysis. *Merrill-Palmer Quarterly, 34*, 261–279.

Authors' Note

The work presented in this article was supported by Grant # NSF 9979968; Early Development of Reading Skills: A Cognitive Neuroscience Approach from the Interagency Educational Research Initiative, with conjoint funding from the National Science Foundation, National Institute of Child Health and Human Development, and U.S. Department of Education. We thank the Houston Independent School District in Houston, Texas, for allowing us to conduct this research and the principals, teachers, and students in the participating schools. Requests for information should be addressed to Patricia G. Mathes.

Conflict of interest statement: Both interventions presented in this research are under contract for publication. The Proactive intervention will be published under a different name. Neither intervention was under contract at the time this research was conducted or during the review process.