KELSEY A. BENSON, NANCY E. MARCHAND-MARTELLA, RONALD C. MARTELLA, and RUSSELL L. KOLTS, Eastern Washington University

Assessing the Effects of the Reading Success Level B Program with Fifth-grade Students at a Title I Elementary School

Abstract: This investigation examined the comparative effects of *Reading Success Level B* on the reading comprehension skills of 78 fifth graders across three general education classrooms. Three student subgroups were formed based on pretest *Scholastic Reading Inventory (SRI)* reading levels including high (*n* = 14), average (*n* = 50), and low (*n* = 14) performance groups. Pre- and posttest reading comprehension measures were collected using the *SRI*. Results showed students in the low performance group demonstrated statistically significant gains in comprehension when compared to students in the average and high performing groups. In addition, students in the average performance group demonstrated statistically significant gains when compared to the high performance group. Risk status did not affect comprehension growth. Implications for future research are discussed.

The acquisition of reading skills is likely the most challenging and complex task students face as they enter the educational system. Not only is reading critically related to many other subject areas including writing, mathematics, spelling, and other language-rich fields of study, strong reading skills are also needed to function adequately in today's literacy-driven society (Van Den Broek & Kremer, 2000). Without these developed reading skills, research has shown that students are more likely to drop out of school (Biancarosa & Snow, 2004; Carnevale, 2001). In addition, poor readers have been known to exhibit higher levels of both internalizing behaviors (such as anxiety and depression) and externalizing behaviors (such as aggression and delinquent behavior) than those students who possess average reading skills (Arnold, Goldston, Walsh, & Reboussin, 2005). With the attainment of reading skills so closely related to students' overall well-being and their functioning within our educational and societal systems, it is no wonder that schools place such great importance on developing this cornerstone of learning (Lyon, 1998).

Despite the considerable importance of reading, many students continue to struggle. According to the National Center for Education Statistics (NCES, 2004), only 31% of fourth graders and 32% of eighth graders performed at or above the proficient level as designated by the National Assessment of Educational Progress (NAEP), with proficient defined as "solid academic performance exhibiting competency over challenging subject matter" (Loomis & Bourque, 2001, p. 6). Furthermore, the largest percentage of students who qualify for special education services have specific learning disabilities in the

area of reading (Culbertson, 1998; Heward, 2006; Meese, 2001).

Due to the widespread reading failure in our nation, the National Reading Panel (NRP) completed a comprehensive research-based report defining five essential areas of effective reading instruction: phonemic awareness (familiarity with individual sounds), phonics (letter-sound associations), fluency (speed, accuracy, and prosody), vocabulary (word knowledge), and text comprehension (understanding of material) (National Institute of Child Health and Human Development [NICHD], 2000). One area of reading development that many students have difficulty with is comprehension (Biancarosa & Snow, 2004; RAND Reading Study Group, 2002). Deficits in this area cause significant problems for the struggling reader because reading comprehension is the "ultimate goal and the essence of reading" (Bos & Vaughn, 2002, p. 182).

Comprehension difficulties often arise when students do not utilize appropriate comprehension strategies while reading. These comprehension strategies provide students with the "tools for understanding the conceptual content of text" (RAND Reading Study Group, 2002, p. 39) so they may become active and purposeful readers (Armbruster, Lehr, & Osborn, 2003). Specifically, these strategies include comprehension monitoring, graphic or semantic organizers, question generation and answering, recognizing story structure, summarization or main idea, and cooperative learning strategies (NICHD, 2000). Research has consistently shown that comprehension strategy instruction has led to gains in reading comprehension performance, even for at-risk students (Biancarosa & Snow, 2004; Dole, Brown, & Trathen, 1996; Mason, 2004).

Comprehension strategies should be taught using a systematic and explicit format (Armbruster et al., 2003; Biancarosa & Snow, 2004; Mason, 2004; RAND Reading Study Group, 2002). Systematic instruction follows a predetermined logical sequence, directly teaching those skills of interest (Armbruster et al.). Explicit instruction "provides a clear explanation of the criterion task, encourages students to pay attention, activates prior knowledge, breaks the task into small steps, provides sufficient practice at every step, and incorporates teacher feedback" (RAND Reading Study Group, 2002, p. 33). Essentially, systematic and explicit instruction involves teacher modeling of new skills in a logical order while providing guided practice exercises and monitoring for student understanding. Only after the concept has been grasped are students allowed to practice the skill independently. When students receive strategy instruction using this instructional format, they are provided with the "how-to information" about the processes involved in comprehending text (Duffy, 2003). While emphasizing the importance of applying strategies during reading, research has consistently shown that systematic and explicit strategy instruction can improve comprehension in struggling readers (Dole et al., 1996; Jitendra, Hoppes, & Xin, 2000; Swanson, 1999).

Despite these findings, few studies have focused on the effectiveness of specific programs designed to teach reading comprehension using an explicit and systematic format. One such program designed to supplement core reading curricula is entitled *Reading Success*. *Reading Success* is a program consisting of Levels *A* (Dixon, Klau, Rosoff, & Conrad, 2002)*, B* (Dixon, Conrad, Salzman, & Klau, 2003), and *C* (Dixon, Conrad, Muti, & Feinberg, 2003) for grades 4, 5, and 6, respectively. Each program level contains 70 to 80 lessons that teach strategies for reading comprehension explicitly, with scaffolding, review, and error correction as central tenets emphasized within the program. Students learn to understand and infer word meanings and figurative language. Comprehension strategies like main idea, summarization, and question answering and generation are

emphasized throughout the program. Additionally, the program provides exposure to a variety of text structures including narrative and expository passages.

To date, one study has been conducted examining the effects of *Reading Success Level A* on the reading comprehension skills of 93 fourth graders (Reed, Marchand-Martella, Martella, & Kolts, in press). Based on pre- and posttest data, substantial reading comprehension growth occurred among all students receiving program instruction. Interestingly, students identified as at-risk for reading failure (students receiving Learning Assistance Program [LAP] tutoring or special education services and/or those with pretest reading comprehension Lexile scores 610L or below) exhibited more comprehension growth when compared to those not considered at-risk. Furthermore, overall reading comprehension growth was not impacted by factors such as gender or which teacher provided instruction. Social validation measures indicated positive teacher perceptions toward the program.

Because no published investigations were found involving the next level of *Reading Success*, the purpose of this study was to assess the affects of *Level B* as the supplemental reading comprehension program specifically designed for fifth graders. Group comparisons similar to those conducted by Reed et al. (in press) were performed. Specifically, the following groups were compared: high, average, and low performance on the *SRI,* and no risk and at-risk (i.e., students who qualified for the school's LAP).

Method

Participants

Students. This investigation included 78 fifth graders, ranging from 10 to 11 years of age, across three general education classrooms. Of these students, 40 (51.3%) were girls and 38 (48.7%) were boys. Approximately 93.5% of

students were Caucasian ($n = 73$), 1.3% were Asian or Pacific Islander ($n = 1$), 1.3% were African American ($n = 1$), 1.3% were Russian ($n = 1$), and 2.6% were Hispanic ($n = 2$).

All students received the *Reading Success Level B* program as a supplement to the regular fifth-grade classroom curriculum. They received reading comprehension instruction in fourth grade using the *Reading Success Level A* program and participated in the study by Reed et al. (in press).

Students were divided into high ($n = 14$, 17.9%), average ($n = 50$, 64.1%), and low ($n = 14$, 17.9%) performance groupings based on their pretest *SRI* reading levels. Those students within one standard deviation of the total group mean were considered average. Those above and below one standard deviation were considered high and low, respectively.

Students also were grouped based on no risk or at-risk status for school failure. Students were considered at risk if they were reading below grade level at less than 105 words read correctly per minute. If they qualified, they were provided LAP services and received instruction involving the program, *Read Naturally* (Ihnot, 1991). Thus, of the 78 students who participated, 54 (69.2%) began the school year with a fifth-grade reading level and read 105 words or more per minute, and 24 (30.8%) qualified for LAP services (considered at risk).

Teachers and graduate student. Three general education teachers provided instruction in *Reading Success Level B.* Teacher A had 28 years of teaching experience over various grade levels (K-6th). She had been trained to use additional Direct Instruction programs including *Reading Mastery* and *Spelling Through Morphographs.* Teacher B had 27 years of teaching experience and had been trained extensively on other Direct Instruction programs including *Spelling Mastery* and *Reading Mastery.*

Additionally, she attended training sessions at the Eugene, OR, Summer Institute in Direct Instruction. For the second half of the school year, Teacher B mentored a student teacher from a local university majoring in elementary education. She did not yet have teaching credentials but had completed all coursework prior to beginning her student teaching experience. After observation and under the supervision of Teacher B, she provided *Reading Success* instruction intermittently. Teacher C had 7 years of teaching experience, with 6 years spent as a special education resource teacher. She also had extensive Direct Instruction training in various programs including *Corrective Reading* and *Spelling Through Morphographs*.

In addition to the teachers involved, a graduate student in school psychology (first author) assisted in re-teaching lessons, administering mastery quiz and test re-takes, and scoring all mastery tests and quizzes. The graduate student had a bachelor's degree in psychology and had completed her first year of graduate study in school psychology when this investigation began. Participation in this study fulfilled her academic research project requirement.

Setting

This study took place in an urban elementary school located in the Pacific Northwest. Out of the 624 students enrolled, approximately 35.6% qualified for free and reduced-price meals. Thus, this school was labeled as a Title I school. As fourth graders during the 2003-04 school year, students were required to take the Washington Assessment of Student Learning (WASL) in the areas of reading and writing. Approximately 94% passed in these two subject areas, respectively. District averages were approximately 83% and 72% for reading and writing, respectively. For these same areas, state averages were approximately 74% and 56%, respectively (Washington State Office of the Superintendent of Public Instruction, 2005).

Targeted Curriculum

Reading Success Level B (Dixon et al., 2003) was implemented in this investigation. *Level B* is a program designed for students with a low to mid-fifth-grade reading level. The program teaches comprehension strategies using explicit and systematic instruction. Teachers demonstrate these strategies (modeling), conduct scaffolded practice so students can apply these strategies (guided practice), and include application of these strategies over time with a variety of examples (independent practice and review). Passages used in *Level B* are longer and have somewhat more complex syntax and higher vocabulary when compared to *Level A*. Like *Level A* (Dixon et al., 2002), *Level B* contains 80 lessons that introduce new concepts and review subsequent concepts learned to that given point. The teaching manual includes all 80 lessons with a 2-in (5.08 cm) left-hand column including a general teacher script with hints on how and when to introduce additional examples or word meanings. The manual also includes answers to quizzes and tests. Students are provided with workbooks so they may follow along with the lessons. These workbooks do not include any quizzes or tests but contain questions for the students to complete independently. A set of black line masters with all quizzes and tests also is provided. Each quiz or test also comes with a black line master score sheet, allowing teachers to keep track of the frequency of missed questions in different content areas. In addition to reviewing anaphora and detail classification from *Level A*, the main concepts introduced in the *Reading Success Level B* program include inference, main idea, fact and opinion, literal comprehension, author's purpose, paraphrase, rewriting passages, word meanings, poetry, figurative language, and several bonus vocabulary words.

Inference. Students are taught to make a variety of inferences based on information from the text. They learn how to infer setting, character

feelings, what a character might have learned, and what might happen next in a story.

Main idea. Students are taught that many comprehension questions relate to the "main idea" of the passage by asking about what the passage is mostly about, what a good title might be, or what might be a good summary or paraphrase of a passage. In finding the main idea, students learn to differentiate between details from the passage, statements that are too general, and good summary statements. In addition, students learn how to write their own main idea statements.

Fact and opinion. Students are taught to discriminate between factual statements and statements of opinion.

Paraphrase. Students are taught to recognize good paraphrasing statements and are taught paraphrasing strategies that can be used when they create their own.

Rewriting passages. Students learn how to rewrite short passages in their own words.

Literal comprehension. Students are provided with ongoing practice of understanding and remembering literal facts and opinions from the text. Questions asking students what may be left out of the passage are a way of emphasizing this fundamental concept.

Poetry. Students are taught to read and understand poetry and are provided instruction on poem structure, including concepts such as stanza and rhyming pattern. Discussion of poem symbolism is provided and students are taught to paraphrase and summarize poetry as well.

Word meanings. Students are taught to understand vocabulary challenges by using the context of the passage to determine meaning.

Author's purpose. Students learn how to identify some of the general purposes of writing (informing, entertaining, or persuading). Students eventually are taught to incorporate the author's purpose into their main idea statements.

Figurative language. Students learn about many forms of figurative language and how to identify and interpret it.

Bonus terminology. Every five lessons, a new "bonus term" is introduced relating directly to reading comprehension. Students are taught the spelling, meaning, derivation, and current usage of the word. The bonus terms for *Level B* are: anaphora, literal, inference, except, metaphor, alliteration, idiom, biography, pseudonym, personification, onomatopoeia, blank verse, plot, symbol, abstract, concrete, and iambic.

Quizzes/tests. Mastery quizzes and tests are built within the framework of the program. Each quiz or test thoroughly and specifically targets concepts learned in previous lessons, with most containing additional bonus questions on terminology. Quizzes are administered every five lessons and usually consist of five to 10 multiple-choice, short answer, or fill-in-the-blank questions. Every 20 lessons, students are given a comprehensive mastery test consisting of 20-25 questions. *Reading Success Level B* contains no placement testing, and all students receive class-wide instruction, beginning with Lesson 1 and ending with Lesson 80.

Additional Curricula

Comprehension instruction occurred in addition to the school's motivational reading program, *Scholastic Reading Counts!* (Scholastic, 2002). *Scholastic Reading Counts!* uses student Lexile scores (Stenner, Smith, Horabin, & Smith, 1988) to determine appropriate grade level books. The students read books and take computerized comprehension quizzes about the books they have selected. Students must pass with 80% accuracy to move on to new books at higher levels. If students do not pass

the comprehension quizzes, they have three opportunities to achieve 80% accuracy. Alternate forms are provided for each quiz to prevent practice effects.

In addition to *Scholastic Reading Counts!,* those fifth graders receiving *Reading Success* instruction also received spelling instruction using the Direct Instruction program, *Spelling Through Morphographs* (Dixon & Engelmann, 2001). *Spelling Through Morphographs* is designed to systematically teach students to break down words into meaningful components with recognizable patterns, based on the principle that identifiable patterns within the English language are useful in spelling words correctly. Students were skill grouped for this program that emphasizes teacher modeling, choral responding, and error correction. Instruction occurred three to four times per week. Mastery tests were given every 10 lessons; review tests were given every 20 lessons.

Additionally, the 24 students who were identified as at-risk readers participated in *Read Naturally* as part of the LAP. *Read Naturally* combines teacher modeling, repeated reading, and progress monitoring to improve reading fluency. It was originally adapted for Title I, English language learners (ELL), and special education populations.

Dependent Variables and Measures

Students were assessed before and after implementation of the program using the *Scholastic Reading Inventory (SRI)* (Scholastic, 1999). In addition, student progress on within-program assessments (tests and quizzes) was recorded. After the program was completed, a social validation questionnaire was given to each teacher to evaluate their opinion of the program.

SRI. Students were assessed before and after the *Reading Success Level B* program using the *SRI.* The *SRI* is a computer-adaptive assessment used to determine how well the stu-

dents read and comprehend literature and expository text at varying levels. The *SRI* focuses on comprehension skills including identifying details in a passage, identifying cause-and-effect relationships and sequencing of events, drawing conclusions, and making comparisons and generalizations. The *SRI* includes authentic text taken from magazines, newspapers, textbooks, and fiction. Depending upon student responses, the computerized assessment proceeds with either easier or more difficult questions. After quiz completion, the *SRI* gives each student a Lexile score, which provides an ordinal scale for measuring student performance within a range of Beginning Reader (BR) to 1700+. After obtaining a Lexile score, students are given a reading level range by adding 50 and subtracting 100 from their respective scores (e.g., *SRI* Lexile = 1000; reading level range = 900 – 1050). These scores allow a student's reading level to be matched with an appropriate text at a corresponding reading level. Although not precisely matched to grade levels, Lexile scores between 750 and 950 are a rough estimate of average fifth-grade performance (K12 Curriculum, 2006).

During this investigation, the Lexile score was used as a measure of progress in reading comprehension. In addition to Lexile scores, *SRI* percentile ranks and stanines were calculated before and after instruction was provided.

Within-program assessments. Student progress on quizzes and tests was monitored and the number of re-takes was recorded for those students not passing the 80% mastery criterion on their first attempt. There were 16 possible re-takes throughout the duration of the program.

Social validation measure. A survey consisting of 10 questions was administered to teachers after all 80 lessons of the program were completed. This survey was developed by Reed et al. (in press) and addressed issues regarding the *Reading Success* program and its implementation. Questions 1, 3, and 8 asked the teach-

ers to provide a rating of the program on a 5-point Likert scale (1 = very poor to 5 = excellent). Question 1 asked about the adequacy of the training for the program. Question 3 asked if adequate support was received while implementing the program. Question 8 asked for a rating of how easy it was to implement the program. Questions 2, 4, 5, 6, 7, 9, and 10 required short answers and comments on how the program could be improved, if differences were seen in student comprehension, and whether or not the program may be implemented in the future.

Design and Procedures

A pre-experimental research design (Martella, Nelson, & Marchand-Martella, 1999) was used during this study for evaluation purposes. All teachers and the graduate student received training before program implementation. Dependent measures were collected before and after the *Reading Success* program was implemented. In addition to the instructional format of the program, regular error correction, concept re-teaching, and re-take procedures of quizzes and tests were followed.

Teacher and research assistant training. A representative from Classical Learning Universe conducted a 2-hr training session at the school for the teachers and graduate student involved in implementing the curriculum. He had a master's degree in special education, was a current special education teacher, and served as a national Direct Instruction consultant. Training consisted of explanation of the different skills learned in the program and the reasoning behind the explicit instructional principles and track sequencing. In addition, testing structure and program materials were explained. Examples of how to provide explicit instruction and procedures for proper error correction also were given.

Reading Success Level B instruction. Regardless of skill level, whole class program instruction took place in three general education class-

rooms approximately three to five times per week unless holidays, conferences, or other school-related activities occurred. Lessons were typically taught in the morning and varied between 15 and 55 min, depending on content. The format of instruction followed the specified outline of the program. Students finished workbook sections according to program guidelines; teachers periodically collected these to monitor their completion. Teachers used choral and individual responding throughout each lesson. Modeling and guided practice of the concepts was used throughout each lesson in accordance with the teaching manual script and lesson design. Throughout lessons, all three teachers provided necessary background knowledge, additional examples, or explanations of new vocabulary terms when needed.

As emphasized in the *Reading Success* program, error correction procedures followed a model, lead, test, re-test approach to error correcting. First, teachers modeled the use of a particular concept explicitly. Next, teachers provided guided practice while students attempted to use the strategy. After sufficient practice, teachers allowed students to practice independently, proving they had mastered the concept or strategy. As needed, review and subsequent guided practice were given.

An example of teaching the concept of main idea from Lesson 11 follows. The teacher might begin by saying, "Class, we will be continuing to learn how to choose a good main idea statement today. Remember, finding a main idea has five important steps. The five steps to finding a main idea are… Get ready." (Point or clap for choral responding). The class would then respond by repeating the five steps to finding the main idea of a passage ([a] decide who or what is talked about the most, [b] write a list of details, [c] classify the details, [d] fill in the main idea boxes, and [e] write a good main idea statement). Class responses could be given through choral or individual responding, depending on teacher

preference. If the class or an individual seemed unclear about any of the steps, the teacher would go over the steps again. For instance, if the class missed the third step (i.e., classify the details), the teacher may say, "Stop. My turn. Remember the five steps to finding the main idea are: decide who or what the passage is about, write a list of details, classify the details (stated with emphasis), fill in the main idea boxes, and write a good main idea statement. When finding main idea, first you decide who or what the passage is about, write a list of details and then you … Everybody, get ready." (Point or clap for choral responding). The class would then respond, "Classify the details." The teacher would continue by giving positive praise statements such as, "Great job identifying the third step to finding the main idea of a passage, classify the details," or by following error correction procedures until the individual or class responded correctly. As students completed work through each lesson, the teacher would walk around and monitor for understanding. If students continued to struggle with specific concepts, the teacher would stop the class and go over the steps again.

As part of the program, teachers administered quizzes and tests every five lessons to determine if students had mastered the concepts introduced in previous lessons. The graduate student scored each quiz or test and recorded each student's score on the score sheets included with the program. Each test or quiz was scored as the number correct out of 100 and reported as a percentage correct. Bonus questions were included and reflected in each student's total score. As recommended by the *Reading Success* program, if more than 25% of the class missed a concept on a quiz or test, the concept was to be reviewed and re-taught to the whole class before subsequent lessons convened. However, because of the variety and volume of missed questions, teachers re-taught only if more than half of the students missed concepts resulting in their failure to achieve 80% mastery. Lessons and subsequent

quizzes or tests were re-taught to the entire class approximately five times across classrooms throughout the duration of the program. If less than half of the students missed a concept and failed to achieve 80% mastery, re-teaching of the concept was conducted by the graduate student. After concept re-teaching, the quiz or test was administered again as a re-take. If students did not reach mastery criteria after this attempt, they moved on in the program with the rest of their class. Re-takes generally occurred once or twice per week or as needed.

Program Fidelity

To ensure instructional fidelity across lessons within *Reading Success Level B*, each teacher was observed on three separate occasions. The graduate student conducted the observations. A previously developed instructional fidelity checklist (Reed et al., in press) was used and included the basic elements emphasized within the program. The teachers were provided with instruction on each point covered on the checklist prior to observation. Teachers were rated from 0 to 5 (0 = does not cover point at all during lesson to 5 = covers point well during lesson) on five criteria. These criteria were: (a) Teacher follows format outlined by *Reading Success Level B* program; (b) Teacher often uses specific praise statements and provides immediate feedback; (c) Teacher monitors student responses frequently during the lesson; (d) Teacher re-teaches either part or all of a lesson(s) when needed and provides alternative or additional explanation(s) when needed; and (e) Teacher uses proper error correction procedures established by the *Reading Success Level B* program. In addition, each major criterion was composed of detailed sub-criteria for clarification. For example, under criterion (e), sub-criteria included "teacher models correct response and has student repeat task error when correcting" and "teacher uses proper amount of repetition and practice to establish mastery of a concept when errors have occurred." After observations were conducted,

teachers were given feedback and copies of their completed observation checklist. Individual scores on the fidelity checklist ranged from 3.0 to 5.0 across criteria, teachers, and observation periods. The overall mean obtained was 4.53 (range 4.0 [Teacher A] to 4.89 [Teacher B]). Thus, the program was correctly implemented at a level of 80 to 98%.

Additional graduate students in school psychology familiar with the fidelity criteria observed and completed program fidelity checklists to ensure agreement on the implementation of the program across classrooms. Interobserver agreement on program fidelity was checked during three observation sessions (one for each teacher). Across observers, teachers, and observation periods, interobserver agreement was calculated with an overall mean of 92.7% (range 92 to 94%).

Statistical Analyses and Comparisons

Comparisons of reading comprehension were made on four variables: performance grouping (i.e., high, average, low), risk status (no risk and at-risk), gender, and teacher. The number of re-takes per student also was explored to evaluate the correlation between re-take count and overall performance.

SRI. Measures of growth were compared using the Wilcoxen signed-ranks test to check for significant differences between pre- and posttest *SRI* Lexile scores, percentile ranks, and stanine scores. *SRI* comparisons also were made among performance groupings (i.e., high, average, low performance; no risk and at-risk). Growth comparisons across groupings were made for *SRI* Lexile scores, percentile ranks, and stanine scores using a Kruskal-Wallis test. These comparisons helped to determine if the program was more beneficial for certain groups of students.

Performance differences in terms of *SRI* Lexile scores, percentile ranks, and stanine scores also were examined among Teachers A,

B, and C. These comparisons were done using a Kruskal-Wallis test to see if scores were impacted by which teacher provided program instruction. Similar comparisons were made by gender using a Mann-Whitney U test.

Within-program assessments. Each student's total number of quiz or test re-takes was recorded and then examined with regard to *SRI* Lexile scores, percentile ranks, and stanine scores using correlational analyses. The average number of student re-takes also was compared across achievement groupings and teachers using a one-way ANOVA. An independent samples t test compared re-take numbers by gender and risk status.

Results

Pretest, posttest, and measures of growth on *SRI* scores (Lexiles, percentile ranks, and stanines) were compared in terms of growth, performance grouping, risk status, gender, and teacher. Within program assessments, correlational analyses, and social validation measures also were obtained. Group means and standard deviations across *SRI* pretest, posttest, and change measures can be found in Table 1. No statistically significant differences were found based on gender and teacher differences on pretest scores, posttest scores, mean growth from pre- to posttest measures, and number of retakes. Thus, no further analyses were conducted for these two categories.

Program Outcomes

Pre- to posttest growth. A Wilcoxen signed-ranks test revealed statistically significant differences across all *SRI* growth measures. More students showed positive ($n = 61$) than negative ($n = 17$) ranks, $z = -5.750$, $p < .001$ on pre- to posttest *SRI* Lexile growth. In terms of percentile rank changes, students showed more positive ($n = 57$) than negative ($n = 14$) or tied ($n = 7$) ranks, $z = -5.525$, $p < .001$. Additionally, more children showed positive ($n = 47$) than negative ($n = 8$) or tied ($n = 23$)

ranks on stanine changes, $z = -5.283$, $p < .001$. These results indicate that most students made positive gains in terms of their *SRI* Lexile, percentile rank, and stanine growth over the course of the school year.

Performance grouping. A Kruskal-Wallis test was used to determine if statistically significant differences on *SRI* Lexile scores, percentile ranks, and stanine scores from pre- to posttest were seen between student performance groups defined as high, average, and low. Results indicated statistically significant differences in change among the high performing group ($n = 14$; *SRI* Lexile: M rank $= 58.43$; percentile rank: M rank $= 61.29$; stanine: M rank $= 53.96$), the average performing group ($n = 50$; *SRI* Lexile: M rank $= 39.30$; percentile rank: M rank $= 39.90$; stanine: M rank $= 41.39$), and the low performing group

Table 1

Pretest, Posttest, and Change Scores

	Total	Performance			Risk Status	
		High	Ave	Low	No Risk	At-Risk
N	78	14	50	14	54	24
SRI Lexile						
Pre	915.79	1231.93	919.36	586.93	1001.96	721.92
(*sd*)	(217.32)	(78.55)	(109.13)	(67.25)	(180.36)	(162.78)
Post	1020.44	1229.07	1026.16	791.36	1094.20	854.46
(*sd*)	(191.01)	(156.25)	(143.08)	(105.08)	(159.27)	(148.91)
Change	103.36	-2.86	104.80	204.43	90.39	132.54
(*sd*)	(126.05)	(115.28)	(115.51)	(85.91)	(124.48)	(127.29)
SRI Percentile Rank						
Pre	64.32	97.29	67.16	21.21	75.24	39.75
(*sd*)	(26.76)	(1.98)	(16.26)	(6.84)	(20.79)	(22.26)
Post	76.54	94.00	79.50	48.50	85.26	56.92
(*sd*)	(20.49)	(10.29)	(15.89)	(14.89)	(13.21)	(19.47)
Change	12.24	-3.29	12.38	27.29	10.06	17.17
(*sd*)	(15.76)	(9.02)	(13.91)	(12.43)	(14.55)	(17.51)
SRI Stanine						
Pre	5.97	8.71	5.96	3.29	6.70	4.33
(*sd*)	(1.87)	(0.47)	(1.05)	(0.61)	(1.60)	(1.31)
Post	6.83	8.43	6.94	4.86	7.46	5.42
(*sd*)	(1.65)	(1.09)	(1.38)	(0.86)	(1.37)	(1.32)
Change	0.86	-0.29	0.98	1.57	0.76	1.09
(*sd*)	(1.13)	(0.91)	(1.06)	(0.65)	(1.16)	(1.02)

($n = 14$; *SRI* Lexile: *M* rank = 21.29; percentile rank: *M* rank = 16.29; stanine: *M* rank = 18.29), *SRI* Lexile $c^2(2) = 18.819$, $p < .001$; percentile rank $c^2(2) = 27.694$, $p < .001$; stanine $c^2(2) = 19.860$, $p < .001$. To further compare the differences between high, average, and low performance groups, post-hoc Mann-Whitney *U* tests were conducted to examine pairwise comparisons between the groups. Significant differences were seen between the high performance and average performance groups for changes in *SRI* Lexile scores [$z(64) = -2.75$, $p < .006$], percentile ranks [$z(64) = -3.70$, $p < .001$], and stanines [$z(64) = -3.53$, $p < .001$], with average performing students exhibiting greater changes in performance than high performing students on all indices. Similarly, significant differences were seen between the average and low performance groups on *SRI* Lexile scores

Table 1, continued

Pretest, Posttest, and Change Scores

	Gender		Teacher		
	Boys	Girls	A	B	C
N	38	40	28	24	26
SRI Lexile					
Pre	928.68	903.55	876.39	922.67	951.88
(*sd*)	(243.94)	(190.98)	(243.98)	(197.41)	(205.38)
Post	1022.58	1018.40	975.32	1075.63	1018.08
(*sd*)	(205.29)	(179.01)	(192.75)	(183.69)	(189.75)
Change	93.89	112.35	98.93	148.79	66.19
(*sd*)	(128.60)	(124.54)	(145.54)	(113.88)	(103.40)
SRI PercentileRank					
Pre	65.58	63.13	58.68	66.00	68.85
(*sd*)	(29.35)	(24.42)	(29.42)	(26.44)	(23.92)
Post	75.74	77.30	71.64	82.29	76.50
(*sd*)	(20.85)	(20.38)	(21.63)	(18.64)	(20.23)
Change	10.16	14.23	12.96	16.38	7.65
(*sd*)	(16.64)	(14.80)	(18.21)	(13.57)	(14.09)
SRI Stanine					
Pre	6.08	5.88	5.68	6.08	6.19
(*sd*)	(2.03)	(1.71)	(2.04)	(1.82)	(1.74)
Post	6.84	6.83	6.46	7.33	6.77
(*sd*)	(1.69)	(1.63)	(1.55)	(1.71)	(1.63)
Change	0.76	0.95	0.79	1.25	0.58
(*sd*)	(1.17)	(1.09)	(1.32)	(0.99)	(0.95)

[$z(64) = -2.91$, $p < .004$], percentile ranks [$z(64) = -3.37$, $p < .001$], and stanines [$z(64) = -1.96$, $p < .05$], with low performing students demonstrating greater changes in performance than did average performing students on all indices.

At-risk status. Additional comparisons were made between those students receiving LAP services (identified as at-risk for reading problems) and those not receiving LAP services (no risk). A Mann-Whitney U test was used to determine if statistically significant differences were seen on pre- to posttest change within *SRI* Lexile scores, percentile ranks, and stanine scores. Compared to their non-risk counterparts ($n = 54$), students receiving LAP services ($n = 24$) did not exhibit statistically significantly differences on measures of *SRI* change. Mean ranks between groups for *SRI* Lexile changes (Mar = 44.94; Mnr = 37.08) were not significantly different, $z = -1.413$, $p < .158$. Mean ranks between groups for percentile rank changes (Mar = 46.31; Mnr = 36.47) also were not significantly different, $z = -1.772$; $p < .076$. In addition, mean ranks between groups for stanine changes (Mar = 43.35; Mnr = 37.79) were not significantly different, $z = -1.043$, $p < .297$. Thus, students considered to be at-risk had similar gains in performance as students not considered at-risk.

Within-program assessments. Among those receiving program instruction, student re-take totals were compared for performance grouping and risk status. Students in the high ($M = 4.07$, $SD = 2.947$), average ($M = 8.04$, $SD = 2.913$), and low ($M = 11.64$, $SD = 2.274$) performance groups exhibited statistically significant differences with respect to their total number of re-takes, $F(2, 75) = 25.284$, $p < .001$. To assess pairwise comparisons among the achievement groups, Tukey post-hoc testing was conducted ($p = .05$). Results indicated that mean re-take counts among the performance groups were significantly different (Mdiff between high and low = 7.57;

Mdiff between high and average = 3.97; Mdiff between low and average = 3.60). Also as expected, an independent samples t test found that at-risk students receiving LAP services exhibited statistically significant differences in their number of re-takes ($M = 10.88$, $SD = 2.771$) when compared to those not receiving LAP ($M = 6.69$, $SD = 3.161$), $t(50.035) = -5.896$, $p < .01$.

Correlations were calculated between total number of re-takes per student and indices of reading function to explore relationships between re-take frequency and outcome variables. Overall, negative correlations were found across pre- and posttest reading measures indicating that as the number of re-takes increased, student performance on each reading measure decreased. These correlations are to be expected because students with lower reading skills would likely need more practice (re-takes) to master the concepts presented. However, it was found that the number of re-

Table 2

Correlational Data on Indices of Reading Functioning

	Pre-Lexile
Number of Re-takes	-.703**
Pre-Lexile	—
Post-Lexile	—
Lexile Change	—
Pre-Percentile Rank	—
Post-Percentile Rank	—
Percentile Rank Change	—
Pre-Stanine	—
Post-Stanine	—

takes was unrelated to measures of change between pre- and posttest *SRI* Lexile scores, percentile ranks, and stanine scores. Table 2 shows correlational data for all reading indices.

Social validity. On the social validation survey, teachers rated the program training average (3.0) to good (4.0) on a 5-point Likert scale, where 1 was very poor, 3 was average, and 5 was excellent ($M = 3.33$). Similarly, teachers rated the support they received during the implementation of the program as average ($M = 3.33$). Teachers rated the program as being average to implement on a daily basis ($M = 3.00$, where 1 was very difficult and 5 was very easy). Finally, teachers were asked open-ended questions regarding the *Reading Success* program. Teachers reported liking the strategies provided by the program and the consistency and repetition that it offered. However, teachers also expressed that concepts were "fuzzy" at times and, although there were many practice exercises, one teacher noted that there was not very much instruction (scripted) included in the program. Teachers stated the quizzes were difficult (requiring multiple retakes for many students) and that some of the stories or poems included seemed "challenging" and "sophisticated" for a fifth-grade audience. For these reasons, teachers questioned their use of *Level B* in the future.

Discussion

This study examined the effectiveness of *Reading Success Level B* (Dixon et al., 2003) on the reading performance of fifth-grade general education students. Results of the statistical analyses indicated that students receiving *Reading Success Level B* instruction showed substantial reading comprehension growth despite gender or which teacher presented instruction. With no differences observed, it can be concluded that the program's effectiveness was not based on the teacher's style or the gender

Table 2, continued

Correlational Data on Indices of Reading Functioning

Post-Lexile	Lexile Change	Pre-Percentile Rank	Post-Percentile Rank	Percentile Rank Change	Pre-Stanine	Post-Stanine	Stanine Change
-.714**	.130	-.673**	-.715**	.214	-.700**	-.708	.124
.819**	-.490**	.980**	.797**	-.629**	.982**	.784**	-.481**
—	.096	.796**	.954**	-.111	.817**	.967**	.062
—	—	-.492**	.062	.920**	-.462**	.106	.921**
—	—	—	.810**	-.645**	.971**	.784**	-.463**
—	—	—	—	-.075	.799**	.962**	.083
—	—	—	—	—	.611**	-.080	.897**
—	—	—	—	—	—	.802**	-.484**
—	—	—	—	—	—	—	.134

*Correlation is significant at the 0.05 level (2-tailed).   ** Correlation is significant at the 0.01 level (2-tailed).

of the student. In addition, students' risk status for reading problems at the beginning of the year did not impact their level of growth on reading comprehension measures. No statistically significant differences were found between at-risk students (those receiving LAP services due to below grade level fluency performance) and their no-risk counterparts in terms of reading comprehension growth. However, student performance groupings (also based on pretest reading comprehension scores) indicated statistically significant differences in the amount of reading comprehension growth accomplished by the students. Post-hoc testing revealed that students in the average performance group exhibited significantly more growth than those in the high performance group. However, students in the low performance group exhibited significantly more comprehension growth than those in the average performance group. These findings suggest that the lowest performing students made the greatest gains in reading comprehension growth, indicating the program may have been most effective for students struggling with reading. However, these students also took significantly more quiz/test re-takes and therefore received the most instruction when compared to the average and high achieving performers.

Other results were observed in this study regarding assessments within the *Reading Success* program. Correlational analyses indicated that there were strong negative relationships between the number of student re-takes required and pre- and posttest measures of reading comprehension. This means that as the number of student re-takes increased (indicating student difficulty achieving mastery criterion on quizzes/tests), reading comprehension scores decreased. Therefore, students who took more re-takes (and thus required additional instruction) had lower overall scores when compared to their peers with less re-takes. However, through the course of this investigation and as discussed earlier, students with the lowest pretest scores made greater reading comprehension progress than their higher achieving peers.

The results of this study may add some important information to the existing research base related to reading comprehension programs. Although the comprehension strategies taught in the *Reading Success* program have been research validated, very few studies have examined the effects of specific explicit and systematic programs designed to increase reading comprehension in upper elementary students. Reed et al. (in press) found that students with lower beginning reading skills exhibited greater growth after receiving reading comprehension instruction from an explicit and systematic program (specifically *Reading Success Level A*). Similarly, students in this study showed comparable patterns of growth. Taken together, this evidence may suggest that students with lower reading skill levels are most likely to benefit from a highly structured, systematic, and explicit reading comprehension program. Although not specifically focused on the *Reading Success* program, other research has shown that struggling readers can benefit from systematic and explicit reading comprehension instruction (Dole et al., 1996; Jitendra et al., 2000; RAND Reading Study Group, 2002; Swanson, 1999).

Despite the information provided by this study, there are also several limitations. First, the study did not utilize a randomized experimental design. Second, all students in this study received reading comprehension instruction in fourth grade using the *Reading Success Level A* program. Without a doubt, this program served to improve student reading comprehension (Reed et al., in press). However, it is possible that because the skills taught in the *Level A* curriculum were very similar to those in *Level B*, true change measures might be more accurately represented by looking at fourth-grade improvements, rather than fifth. Also, because all students had the *Reading Success Level A* program in fourth grade, the results obtained from this study may not gen-

eralize to other fifth-grade populations. For comparison purposes, future research on this program may want to include groups of students who did not receive *Level A* instruction before *Level B*.

Third, this study took place in a school that has consistently scored high on the WASL test in the areas of reading and writing. The successful reading and writing achievement of this particular school also may make these results more difficult to generalize to other fifth-grade populations. Future studies may want to focus implementation of reading comprehension programs in schools that do not have such demonstrated success, to see if similar results are found.

Fourth, although students received *Reading Success* instruction in both *Level A* and *Level B*, no long-term measure was obtained regarding the maintenance of these reading comprehension skills. Due to the mastery component of this program, it is likely that skills would be maintained; however, no data exists to support this hypothesis. Future research should address long-term maintenance of such comprehension skills.

Finally, the level of social validity indicated by the teacher's rating of the *Reading Success* program may be seen as a limitation of this study. Due to the difficulty of the program content and the necessity of re-teaching (and subsequent quiz/test re-takes), teachers expressed they probably would not use *Level B* as a program for reading comprehension in the future. Pressley and El-Dinary (1997) indicate the importance of teachers "buying into" instructional programs, which is based largely upon their perceptions of student success. According to the teachers in this study, perception of student success was relatively low despite demonstrated gains. Future research should continue to take the impact of social validity into consideration when evaluating programs.

Overall, this study demonstrated successful implementation of *Reading Success Level B*, particularly for those students beginning the school year with lower reading skills. With reading comprehension a significant area of concern among students in our nation, educators need to be aware of the importance of good instruction, as it "is the most powerful means of developing proficient comprehenders and preventing reading comprehension problems" (RAND Reading Study Group, 2002, p. 29).

References

Armbruster, B. B., Lehr, F., & Osborn, J. (2003). *Put reading first: The research building blocks of reading instruction (Kindergarten through Grade 3)* (2nd ed.). Washington, DC: Partnership for Reading, Center for the Improvement of Early Reading Achievement.

Arnold, E. M., Goldston, D. B., Walsh, A. K., & Reboussin, B. A. (2005). Severity of emotional and behavioral problems among poor and typical readers. *Journal of Abnormal Child Psychology, 33(2)*, 205-218.

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

Bos, C. S., & Vaughn S. (2002). *Strategies for teaching students with learning and behavior problems* (5th ed.). Boston: Allyn and Bacon.

Carnevale, A. P. (2001). *Help wanted…college required.* Washington, DC: Educational Testing Service.

Culbertson, J. L. (1998). Learning disabilities. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of Child Psychopathology* (3rd ed.) (pp. 117-156). New York: Plenum.

Dixon, B., Conrad, L., Muti, K., & Feinberg, L. (2003). *Reading Success: Level C.* Boston: Classical Learning Universe.

Dixon, B., Conrad, L., Salzmann, K., & Klau, K. (2003). *Reading Success: Level B.* Boston: Classical Learning Universe.

Dixon, B., Klau, K., Rosoff, A., & Conrad, L. (2002). *Reading Success: Level A.* Boston: Classical Learning Universe.

Dixon, R., & Engelmann, S. (2001). *Spelling Through Morphographs.* Columbus, OH: SRA.

Dole, J. A., Brown, K. J., & Trathen, W. (1996). The effects of strategy instruction on the comprehen-

sion performance of at-risk students. *Reading Research Quarterly, 31*(1), 62-88.

Duffy, G. (Ed.). (2003). *Improving comprehension – 10 research-based principles.* Washington, DC: National Education Association.

Heward, W. L. (2006). *Exceptional children: An introduction to special education* (8th ed.). Columbus, OH: Merrill.

Ihnot, C. (1991). *Read Naturally.* St. Paul, MN: Read Naturally Incorporated. Retrieved March 15, 2005, from http://www.readnaturally.com

Jitendra, A. K., Hoppes, M. K., & Xin, Y. P. (2000). Enhancing main idea comprehension for students with learning problems: The role of a summarization strategy and self-monitoring instruction. *Journal of Special Education, 34*(3), 127-139.

K12 Curriculum. (2006). *Language Arts 5.* Retrieved February 2, 2006, from http://www.k12.com/curriculum/subjects/language/language_5.html

Loomis, S. C., & Bourque, M. L. (Eds). (2001). *National Assessment of Educational Progress Achievement Levels, 1992–1998 for Reading.* Washington, DC: National Assessment Governing Board.

Lyon, R. G. (1998). Why reading is not a natural process. *Educational Leadership, 55*(6), 14-19.

Martella R. C., Nelson, J. R., & Marchand-Martella, N. E. (1999). *Research methods: Learning to become a critical research consumer.* Boston: Allyn and Bacon.

Mason, L. (2004). Explicit self-regulated strategy development versus reciprocal questioning: Effects on expository reading comprehension among struggling readers. *Journal of Educational Psychology, 96*(2), 283-296.

Meese, R. L. (2001). *Teaching learners with mild disabilities: Integrating research and practice* (2nd ed.). Belmont, CA: Wadsworth/Thomson Learning.

National Center for Education Statistics. (2004). *The Nation's Report Card.* Retrieved February 25, 2005, from http://www.nces.ed.gov/nationsreportcard/reading

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: Reports of the subgroups* (NIH Publication NO. 00-4754). Washington, DC: U.S. Government Printing Office.

Pressley, M., & El-Dinary, P. B. (1997). What we know about translating comprehension-strategies instruction research into practice. *Journal of Learning Disabilities, 30*(5), 486-488.

RAND Reading Study Group. (2002). *Reading for understanding: Toward an R & D program in reading comprehension.* Santa Monica, CA: RAND. Retrieved March 25, 2005, from http://www.rand.org/multi/achievementforall/reading/readreport.html

Reed, J., Marchand-Martella, N., Martella, R., & Kolts, R. (in press). Assessing the effects of the *Reading Success Level A* program with fourth-grade students at a Title I elementary school. *Education and Treatment of Children.*

Scholastic. (1999). *Scholastic Reading Inventory.* New York: Scholastic.

Scholastic. (2002). *Scholastic Reading Counts!.* Retrieved March 15, 2005, from http://www.src.scholastic.com

Stenner, A. J., Smith, M., Horabin, I., & Smith, D. (1988). *User's manual for the lexile test of reading comprehension.* Durham, NC: MetaMetrics.

Swanson, H. L. (1999). Reading research for students with LD: A meta-analysis of intervention outcomes. *Journal of Learning Disabilities, 32*(6), 504-532.

Van Den Broek, P., & Kremer, K. (2000). The mind in action: What it means to comprehend during reading. In B. M. Taylor, M. F. Graves, & P. Van Den Broek (Eds.), *Reading for meaning: Fostering comprehension in the middle grades* (pp. 4-23). New York: College Teachers Press.

Washington State Office of the Superintendent of Public Instruction. (2005). *School Report Card.* Retrieved February 15, 2005, from http://reportcard.ospi.k12.wa.us

Author Note