JOURNAL OF EDUCATION FOR STUDENTS PLACED AT RISK, *7*(2), 141–166
Copyright © 2002, Lawrence Erlbaum Associates, Inc.

Increasing the Reading Achievement of At-Risk Children Through Direct Instruction: Evaluation of the Rodeo Institute for Teacher Excellence (RITE)

Coleen D. Carlson and David J. Francis

Texas Institute for Measurement, Evaluation, and Statistics
University of Houston

This article describes an evaluation of a program, the Rodeo Institute for Teacher Excellence (RITE), which addresses at-risk students' failure to develop reading skills. The evaluation included all Grade K–2 students participating in the program and in comparison schools selected to serve as a control group. Results indicated that the RITE program was successful at increasing the reading abilities of students in at-risk schools. Children who began the RITE program early and who spent more years in the program outperformed all other students. Intervention with teachers was related to improvement in observed teaching skills (behavior management and teacher corrections), and successful implementation of programmatic teaching techniques was related to student performance. These findings close the trainer–teacher–student feedback loop by showing that teacher behavior relates to student performance.

Failure to develop basic reading abilities during the first few years of school has been shown to be related to a number of academic, economic, and socioemotional difficulties (Lipson & Wixson, 1997; Pressley & Hampston, 1998; Snider & Tarver, 1987; Wharton-McDonald, Pressley, & Hampston, 1998). Juel (1988) reported that approximately 88% of first-grade students whose performance scores were in the lower quartile in reading comprehension remained at performance levels below the 50th percentile through fourth grade. Others have reported similar findings in that students who have been poor readers in the early elementary years remain poor readers throughout school; the problem intensifies with each new year. Concern over early reading and the prevention of early reading problems has resulted in two national research reviews in the last 3 years, one commissioned by the National Research Council (Snow, Burns, & Griffin, 1998), and the other by the

Requests for reprints should be sent to Coleen D. Carlson, Texas Institute for Measurement, Evaluation, and Statistics, 126 Heyne Building, Department of Psychology, University of Houston, Houston, TX 77204–5022. E-mail: coleen.carlson@times.uh.edu

U.S. Congress, coordinated by the National Institute for Child Health and Human Development (National Reading Panel Report, 2000).

Central to the acquisition of early reading ability is the speed and comprehensiveness with which children learn the process of decoding. For many students, learning the alphabetic principle is easy. Many students enter school having already experienced a variety of literacy-related activities, and many already have at least some knowledge of letters and sound–symbol (letter–sound) correspondence. However, other students have had significantly fewer literacy-related opportunities prior to those first provided within the school setting and enter school with little to no knowledge of the alphabetic principle. At least in part due to these more limited experiences, these same children have been shown to be those who are less likely to develop automatic decoding skills (Adams, 1990). Although being at risk for reading difficulties is not a circumstance limited to students from lower economic strata or to those living in urban settings, the prevalence of lower reading performance levels and less developed reading abilities for students from this background in urban settings tends to be significantly higher.

For at-risk students, the effectiveness of the reading instruction they receive in the early school years is of utmost importance. There has been considerable debate over the past decade about what constitutes the most effective beginning reading instruction. In the empirical literature of reading instruction, explicit decoding instruction has increasingly been cited as a more effective instructional approach (Stahl, McKenna, & Pagnucco, 1994) than more implicit methods. Studies have shown that the standardized tests scores of students participating in programs that explicitly teach phonemic awareness, phonics, and letter–sound associations increase or are at higher levels than those of students in other types of programs (Adams, 1990; Pflaum, Walberg, Karegaines, & Rahsher, 1980). Other studies of tutoring intervention have shown that the more successful interventions with at-risk students include higher occurrences of modeling, word study practices, and more time spent practicing skills (Juel, 1996; Leslie & Allen, 1999).

Program effectiveness is only partially attributable to the content of the instructional program itself. The teacher's implementation of the program is equally important to the successful development of reading skills in at-risk children. Because many teachers are not adequately prepared for the task of teaching reading to at-risk children, teacher training and professional development are also important factors in decreasing the number of students who fail to develop basic reading skills (Brady & Moats, 1997).

The Rodeo Institute for Teacher Excellence (RITE) was designed to provide severely at-risk Grade K–2 students with explicit instruction in phonemic awareness and decoding through a consistent curricula, adequate materials, and skilled teachers. The RITE program was modeled after a successful program implemented in one elementary school within the district that saw dramatic skill gains in its students. At the core of the RITE program is phonics-based instruction and an empha-

sis on professional development. The foundation of this program rests in the reading mastery (RM) curriculum (Englemann & Bruner, 1995). Using these instructional materials, the RITE program strives to strengthen teachers' skills in reading instruction through intensive teacher training and year-long support provided by trainers who work within the schools and consult with teachers, providing feedback on program planning and implementation. During the summer, each teacher attends in-depth, hands-on phonics instruction training using the RM materials. During the school year, each school is assigned a master trainer, who provides daily on-site support and holds monthly meetings with all teachers to discuss issues and concerns in an open forum.

In addition to the teaching skills directly related to the RM curricula, the RITE program also strives to provide teachers with strong classroom management techniques. Many of the teacher skills focused on in the RITE program are similar to those that have been presented in literature on the best teaching practices. Specifically, the RITE program emphasizes the consistent and complete use of the "model–test–retest" correction technique from the RM curriculum. This technique directs students' attention to mistakes, provides them with a model of the correct response, tests their knowledge after the modeling has occurred, and continues this practice until students are firm in their knowledge. Furthermore, the program also emphasizes the use of positive reinforcement in the form of teacher praise and verification for correct student responses. Finally, the program emphasizes strong classroom management skills. The ability of the teacher to provide an environment that is safe, consistent, and that allows the student time to focus on the tasks at hand is both an important goal of this program and necessary to its success. Because students receive reading instruction in this program in small, skill-leveled groups, it is imperative that those students not in a reading group remain on task and engaged in active learning activities (e.g., independent reading, listening center activities).

The RITE program completed its fourth year of implementation during the 2000–2001 academic year. Each year, the program has increased in size, beginning in 6 schools during the 1997–1998 school year and ending the 2000–2001 school year in 20 schools. The external evaluation of the RITE program initially focused only on student performance levels and gains. Over time, the evaluation expanded to include observations of teacher performance within the classroom; trainer reports of intervention provided to teachers; and examination of the links between student, teacher, and trainer performance.

PARTICIPANTS

Program Group

The external evaluation included students from all kindergarten, first-, and second-grade classrooms from all years of the RITE program. Third-grade perfor-

mance levels were also examined for those students who had participated in the program at some point during their kindergarten, first-, or second-grade year; and for whom data were available from the district on the state-mandated assessment, the Texas Assessment of Academic Skills (TAAS). Table 1 describes the number of students beginning each year in each grade, as well as the number of students from each grade level who participated in the program across years. Roughly equal numbers of boys (51%) and girls (49%) participated in the RITE program, and the majority of the RITE students were of African American (65%) or Hispanic (28%) descent. The remainder of the participants were White (3%), Asian American (3%), and American Indian (1%).

In the 2000–2001 academic year, 277 teachers from 20 schools participated in the RITE program. Of these teachers, 137 were new to the RITE program, 74 were returning for their second year, 47 for their third year, and 19 for their fourth year. Of the 20 RITE schools, 4 were entering their fourth year in the program, 3 were entering their third year in the program, 6 were entering their second year, and 7 were beginning their first year with the program.

Comparison Group

Twenty comparison schools were selected to serve as a control group for each of the RITE program schools. Each year, as schools were added to the RITE program, comparison schools were selected from the non-RITE schools in the district. These schools were selected based on the degree to which school characteristics matched those of one of the RITE schools. The characteristics examined included the percentage of students who

1. Received free or reduced-price lunch.
2. Belonged to an ethnic minority group.
3. Were limited English proficient.
4. Met the minimum state-mandated reading performance requirement.

All non-RITE schools in the district were compared to the participating RITE schools during the RITE schools' first year of participation in the program. If the percentages in all four categories for a comparison school were within a 10% range of the percentage at a particular RITE school, then that comparison school was placed in a pool of possible matches for that RITE school. Because each RITE school had more than one possible comparison school, the next step was to identify the comparison school that was geographically closest to the RITE school. The comparison school within the pool of possible matches for a RITE school that was geographically closest to that RITE school was selected as that RITE school's match.

The comparison group provided a means for judging the performance of students in the program relative to expectations for similar students who were participating in

TABLE 1
RITE Students by Grade Within and Across Program Years

		RITE Program Students Over Time			
Grade	*Year Entered Program*	*Kindergarten*	*First Grade*	*Second Grade*	*Third Grade*
Kindergarten	1997–1998	413 (1997–1998)	259 (1998-1999)	181 (1999–2000)	137 (2000–2001)
	1998–1999	607 (1998–1999)	420 (1999–2000)	271 (2000–2001)	—
	1999–2000	969 (1999–2000)	611 (2000–2001)	—	—
	2000–2001	1460 (2000–2001)	—	—	—
First grade	1997–1998	—	440 (1997–1998)	268 (1998–1999)	192 (1999–2000)
	1998–1999	—	494 (1998–1999)	252 (1999–2000)	163 (2000–2001)
	1999–2000	—	1045 (1999–2000)	624 (2000–2001)	—
	2000–2001	—	1036 (2000–2001)	—	—
Second grade	1997–1998	—	—	457 (1997–1998)	301 (1998–1999)
	1998–1999	—	—	462 (1998–1999)	314 (1999–2000)
	1999–2000	—	—	986 (1999–2000)	546 (2000–2001)
	2000–2001	—	—	993 (2000–2001)	—

Note. RITE = Rodeo Institute for Teacher Excellence.

other district programs. Although not equivalent to randomizing, it was an attempt to provide a baseline performance standard for children with similar demographic characteristics that were attending schools of similar composition in similar geographic regions of the same district. Insofar as the district has had an active program targeting improved reading performance, it is critical that any outcomes associated with participation in the RITE program be judged relative to outcomes that could have reasonably been expected for these students had their school not participated in the RITE program. Due to the emphasis on improved reading performance at the district level, each of the comparisons schools was required to provide a reading curriculum for kindergarten, first, and second grades. However, within the comparison schools, the curricula across grades was not as standardized as it was in the RITE program, nor was not the same level of support provided to teachers regarding program implementation and implementation skill development.

MEASURES

The skill assessments administered at each grade were chosen to capture the multicomponent nature of academic reading skills in kindergarten and Grades 1 and 2, as well as the central importance of the TAAS examinations for Texas public school children in Grades 3 and beyond. In the first year of the program (1997–1998), all kindergarten students were administered the Word Identification subtest of the Woodcock–Johnson Mastery Test–Revised (Woodcock & Johnson, 1979) in the fall and spring of the school year. In 1998–1999, the assessment plan was modified, and individual assessments of students by trained pyschometricians were eliminated in favor of collecting district-mandated assessments, in an effort to reduce the overall cost of the evaluation program and to reduce the testing burden to students. Unfortunately, because this was the first year of a new assessment program within the district, some problems were experienced with the data collection mechanisms at the district level, specifically in the transfer of data from teacher-administered assessments in the fall and spring to district reporting forms. As a result, no data were available that year for the kindergarten students. In the 1999–2000 and 2000–2001 academic years, all kindergarten RITE and comparison school students' scores from both the winter and spring district-administered screening section of the Texas Primary Reading Inventory (TPRI) were collected. In each of the four program years, all first- and second-grade RITE and comparison school students' Word Reading and Reading Comprehension scores from the district-administered Stanford Achievement Test–Ninth Edition (SAT9) were collected. In the 2000–2001 academic year, Word Reading scale scores from the SAT9 were also available for kindergarten students. For those students who participated in the RITE program at some point during kindergarten, first, or second grade, and who have since reached third grade, the Texas Learning Index from the district-administered TAAS were also collected.

Woodcock Reading Mastery Test–Revised (Woodcock & Johnson, 1979)

In 1997–1998, all RITE and comparison kindergarten students were administered the Word Identification subtest in the fall and spring of the school year. The Word Identification task asks students to read words presented one at a time and consists of a total of 20 items. Internal consistency reliability coefficients for this subtest were .91 and .92, respectively, for the fall and spring administrations.

TPRI

The TPRI is a teacher-administered instrument developed to assist the teacher in identifying students' skills and skill levels and to guide instruction (Texas Education Agency [TEA], 1998). By district mandate, the TPRI was administered to kindergarten students by their classroom teacher in the winter and spring of each school year (beginning in 1998–1999). In kindergarten, the TPRI screen consists of measures of Phonological Awareness—specifically, Letter–Sound Identification and Blending Onset and Rime.

Letter–sound identification. The Letter–Sound Identification section of the TPRI screen consists of 10 letters of which the child must correctly provide the associated sound for 8 to "pass" the screen. The letters on this screen are considered to be 10 of the more difficult letters for children to learn the associated sounds, including: L, O, N, I, R, E, H, W, U, and Y. This portion of the TPRI screen has reliability (coefficient alpha) of .90 and a bivariate correlation with end of Grade 1 reading of .54.

Phonological awareness. In the Phonological Awareness (Blending Onset and Rime) section of the TPRI, the child is presented with isolated pairs of onset and rimes and asked to put the two parts together to make a word. There are eight items on the screen; a score of six out of eight correct is considered passing. This portion of the TPRI screen has reliability (coefficient alpha) of .91 and a bivariate correlation with end of Grade 1 reading of .50.

SAT9

The SAT9 is a norm-referenced standardized test that is designed to measure performance in the areas of reading, spelling, study skills, language, mathematics, science, and social science. The SAT9 is a district-mandated achievement test, and, for this

evaluation, students' scores on the Word Reading and Reading Comprehension subtests were collected from the district research and accountability department.

TAAS

The TAAS is a state-mandated, criterion-referenced assessment given to all third-grade students in the spring of each academic year (TEA, 1990). In this evaluation, the Texas Learning Index (TLI) scores from the reading portion of the TAAS were collected for all RITE and comparison school children once the child reached third grade. The TLI score is a modified *t* score of the student's raw score. Specifically, the TLI is a *t* score that is anchored at the exit level passing standard rather than the mean of the distribution. A TLI score of 70 or above is considered passing, or indicates that the student has met the minimum standards for that grade level.

Teacher Observations

During the 2000–2001 school year, all kindergarten, first-, and second-grade classrooms were observed in each of the RITE program schools. Each classroom was observed at two time points (the fall of 2000 and the spring of 2001) and on two different days at each time point. During each observation session, a trained research assistant observed the teaching of a lesson for a period of approximately 20 to 25 min. Therefore, each teacher was observed for an average of approximately 80 to 100 min over the course of the year. A different observer conducted each of the two observations at a given time point in a particular classroom, and at each time point, one observation was conducted during instruction with the highest reading group, and the other with the lowest reading group in each classroom. Therefore, the possibility that teacher behaviors and techniques are simply a result of the ability level of the group being taught is minimized.

During the observation, two areas were alternately the focus of observations: teacher corrections (i.e., whether the teacher, in response to student errors, provided the group or individual with corrections that followed the RITE program model–test–retest paradigm), and teacher responses to student responses (i.e., whether praise, verification of the student response, or both, was provided; or whether no teacher response was provided). During an observation, the unit of focus in the classroom observations was considered to be a response. Therefore, information was recorded for each and every response requested by a teacher or provided by a child. The time that the observer spent in the classroom was divided into six segments. The observer rated each of the two aforementioned categories during three different segments. Therefore, during one observation period, each category was focused on three times. Hence, across the two observation sessions in

a given time period within the year, each category was observed six times. The decision to focus on these two techniques was based on the importance that the RITE program placed on these key programmatic instruction techniques.

In addition to the two programmatic instructional techniques described earlier, classroom management is also an important aspect of the RM program. Consequently, observers also recorded the number of behavioral interruptions that occurred that caused a break in the flow of the reading group and the number of children outside of the reading group who, during the session being observed, were not engaged in independent activities.

Trainer Support Survey

Each RITE trainer was asked to complete a feedback survey on teachers in their schools. Trainers were asked to rate the level of preintervention problems, postintervention problems, and the general amount of intervention they provided each teacher in four different areas. Of the four areas rated, two focused on general classroom teaching, and two on skills specific to aspects of the RITE program. Specifically, the two general classroom teaching skill items were "classroom management–organization" and "disciplinary techniques–behavior management." The three programmatic teaching skills reported on included "understanding key concepts of the program" and "appropriate use of teacher corrections."

The scale for rating teachers' problems pre- and postintervention was a 4-point Likert-type scale, ranging from 1 (*no problem*) to 4 (*seriously problematic*). Ratings of the level of intervention provided were also reported on a 4-point Likert-type scale, ranging from 1 (*no help was ever provided*) to 4 (*the area was addressed specifically on numerous occasions*). Trainers' ratings of the four areas were averaged into the two teaching skills categories according to the list mentioned earlier (general classroom management and programmatic teaching).

PERFORMANCE ANALYSES

Preliminary Analyses

Prior to performing all analyses, all performance scores were examined for outlying values (e.g., specific children whose scores were much higher or lower than the majority), and the data was then checked to ensure the accuracy of the scores. Next, the number of teachers and children in the RITE and comparison schools were compared to ensure relative equivalence. Each year, attrition analyses were conducted to examine the performance levels of students who left the program school versus those who remained in the program school into the next grade (Little & Rubin, 1987;

Shafer, 1997). No significant differences in performance were found between the attrition and non-attrition groups in any analyses. Furthermore, attrition was not related to student characteristics such as gender and ethnicity. Based on these results, it is reasonable to consider that factors other than student performance and student demographic characteristics are responsible for attrition and that, for the sake of analyses, the data meet the assumptions for missing at random (Little & Rubin).

Within-Grade Analyses

The first set of analyses examined student performance in each grade separately. Because children learn and perform in similar settings (e.g., classrooms) and receive instruction from similar sources (e.g., teachers), scores for all students in the same classroom are not independent of one another. This lack of independence among observations must be taken into account in the analyses and makes the use of conventional analyses problematic. Instead, to account for the non-independence of students' performance, multilevel modeling techniques were employed in all analyses. The two levels in these analyses included the individual and the classroom (or teacher). Inclusion of the second level (teacher) addresses the possible non-independence among scores for children in the same classroom.

All within-grade analyses examined performance differences between comparison and RITE students as a function of the number of years of program experience. Therefore, we can compare the relative performance levels of all first graders and all second graders as a function of the number of years the children have participated in program schools. For kindergarten students, performance was examined as a function of RITE versus comparison only, as all kindergarten students had only 1 year in the program. Table 2 presents the total number of students included in the within-grade analyses for each grade level by number of years in the RITE program.

Kindergarten analyses. The analysis of kindergarten students' performance was conducted for students in the 1997–1998, 1999–2000, and 2000–2001 academic years. During the 1998–1999 school year, no kindergarten assessment data was available. In 1997–1998, kindergarten students were administered the Word Identification subtest of the Woodcock–Johnson Mastery Test–Revised (Woodcock & Johnson, 1979). In the 1999–2000 school year, TPRI data were available for each student; and in 2000–2001, TPRI as well as SAT9 Word Reading data were available for kindergarten students.

Analyses of the 1997–1998 kindergarten data examined group differences in spring performance using fall performance as a covariate, in addition to examining group differences in fall performance, and evidence for differential effectiveness of the fall covariate (i.e., heterogeneity of regression). Results indicated that groups did not differ in the fall, $F(1, 412) = .08$, $p \leq .78$; and there was no evidence for heteroge-

neity of regression. However, after controlling for fall performance levels in Word Identification skills, the RITE kindergarten students' performance levels were significantly higher than those of the comparison students, as shown in the top section of Table 3, $F(2, 412) = 17.42$, $p \leq .0001$. Therefore, RITE kindergarten students showed statistically greater gains in Word Reading skills over the course of the academic year than did comparison students. Although there was a small difference between groups at the pretest, it is important to bear in mind two facts when considering this difference. First and foremost, the difference is not statistically significant, indicating that we cannot reject the possibility that the groups are equivalent at the pretest. Second, the difference is small, such that even if we reject the notion that the groups are equivalent at the pretest, they are not largely different.

Analysis of the pass rates on the TPRI across the kindergarten year included two cohorts of kindergarten students (those from the 1998–1999 school year, as well as those from the 1999–2000 school year). Analyses examined differences in RITE and comparison students pass rates on both the winter and spring TPRI, as well as the relative gain in the number of students attaining passing status over the kindergarten year. From these analyses, we can ascertain not only whether more RITE children are passing the skills sections of the TPRI than comparison children at the end of the kindergarten year, but whether children are more likely to drop their at-risk status as identified by the TPRI screening when they receive kindergarten instruction through the RITE program.

Table 4 presents the pass versus no pass and RITE versus comparison group status for middle and end-of-year TPRI scores. Results of these analyses indicate that there were significant group differences (RITE vs. comparison) in students' TPRI pass rates in both the winter and spring. Specifically, in the middle of the kindergarten year, more RITE kindergarten students were passing the Letter–Sound Identification sections of the TPRI: RITE = 65%, comparison = 42%, $\chi^2 = 32.95$, $p \leq .001$;

TABLE 2
Total Number of Students by Number of Years in the RITE Program Through the 2000–2001 Academic Year

		Number of Years in the RITE Program			
Treatment Group	*Grade Level*	*0*	*1*	*2*	*3*
COMP	Kindergarten	2,105	—	—	—
	First grade	3,924	—	—	—
	Second grade	2,838	—	—	—
RITE	Kindergarten	—	2,842[a]	—	—
	First grade	—	3,015	1,290	—
	Second grade	—	2,898	1,144	452

Note. COMP = comparison; RITE = Rodeo Institute for Teacher Excellence.
[a]Unavailable data in the 1998–1999 school year.

TABLE 3
Kindergarten Woodcock–Johnson and SAT9 Performance Means and Percentiles

	Treatment Group	
Subtest	*RITE*	*Comparison*
Woodcock–Johnson Word Identification Fall (1997–1998)		
M	0.84	0.75
SD	0.22	0.18
Woodcock–Johnson Word Identification Spring (1997–1998)		
M	16.76*	6.90*
SD	1.29	1.21
SAT9 Word Reading (2000–2001)		
M	457.71**	435.66**
SD	54.92	44.52
Below 25th Percentile	13%	22%
Above 50th Percentile	69%	49%

Note. RITE = Rodeo Institute for Teacher Excellence; SAT9 = Stanford Achievement Test–Ninth Edition.

*$p \leq .0001$, after controlling for beginning of the year performance. **$p \leq .0001$.

and there was no significant differences in pass rates on the Phonological Awareness section of the TPRI (i.e., Blending Onset Rimes) between RITE and comparison school students: RITE = 19%, comparison = 15%, $\chi^2 = .44$, $p \leq .51$. At the end of the year, results indicated that more RITE children were passing both the Letter–Sound Identification: RITE = 91%, comparison = 78%, $\chi^2 = 26.32$, $p \leq .001$; and the Phonological Awareness sections of the TPRI screen: RITE = 68%, comparison = 50%, $\chi^2 = 28.32$, $p \leq .001$. Although there were no program group differences in the percentage of students moving from failing to passing from the middle to the end of the year on the Letter–Sound Identification section of the TPRI (RITE = 27%; comparison = 39%), more RITE children were moved from failing to passing status on the Phonological Awareness section of the TPRI over the course of the year than comparison children (RITE = 50%; comparison = 39%).

The final set of kindergarten analyses examined group differences in the SAT9 Word Reading skills for the 2000–2001 kindergarten students. These analyses compared average performance levels across the RITE and comparison groups, as well as the percentage of children performing at or below the 25th percentile, as well as those performing at or above the 50th percentile. Although the comparison of mean scores across groups provides useful information about general levels of performance, examining the distribution of percentile scores within and between groups provides important information about whether the RITE program is reduc-

TABLE 4
Kindergarten TPRI Pass Rates From Middle to End of Year

TPRI Subtest	*Group*	*Middle of the Year*	*End of Year*	
			No Pass	*Pass*
		No pass	8%	27%
Letter–Sound Identification	RITE	Pass	1%	64%
		No pass	19%	39%
	COMP	Pass	3%	39%
		No pass	31%	50%
	RITE	Pass	1%	18%
Phonological Awareness		No pass	46%	39%
	COMP	Pass	4%	11%

Note. COMP = comparison; RITE = Rodeo Institute for Teacher Excellence; TPRI = Texas Primary Reading Inventory.

ing students' risk for low achievement or producing achievement levels that exceed normative expectations. Results indicated significant differences in RITE versus comparison students SAT9 Word Reading levels, such that the RITE kindergarten students were performing at significantly higher average skill levels than comparison students by the end of the kindergarten year (see lower section of Table 3), $F(1, 1{,}459) = 99.47$, $p \leq .0001$. Furthermore, RITE kindergartners were less likely to score below the 25th percentile, and significantly more likely to score above the 50th percentile than comparison students.

Taken together, results of the kindergarten analyses indicate that students in the RITE program show significantly higher levels of phonemic awareness and word reading skills than peers not in the program. Furthermore, in regard to word reading, the RITE kindergarten students are not only performing at higher levels than their nonprogram peers, but are also performing above national norms as indexed by percentile scores on the SAT9 Word Reading subtest.

First- and second-grade analyses. Two types of analyses were conducted in the examination of all first- and second-graders performance; and for each of these types of analyses, SAT9 Word Reading and Reading Comprehension performance were examined separately. The first set of analyses compared average performance levels on the SAT9 skills tests as a function of the number of years of experience the students had in the RITE program. The second set of analyses examined the average percentage of children across these same experience groups who were performing at or below the 25th percentile as well as those performing at or above the 50th percentile. Average performance levels are presented in Table 5 for first and second grades, respectively.

TABLE 5
First and Second Grade End of Year Stanford Achievement Test–Ninth Edition Means and Percentile Performance

	RITE Years			*Comparison Years*		
Subtest	*3*	*2*	*1*	*3*	*2*	*1*
First grade						
Word Reading						
M	—	541.51	516.98	—	509.83	506.79
SD	—	55.48	52.87	—	51.79	51.79
Below 25th percentile	—	8%	20%	—	26%	23%
Above 50th percentile	—	73%	52%	—	48%	47%
Reading Comprehension						
M	—	549.86	533.12	—	518.31	516.31
SD	—	49.37	47.97	—	47.71	45.18
Below 25th percentile	—	11%	19%	—	26%	24%
Above 50th percentile	—	78%	60%	—	55%	52%
Second grade						
Word Reading						
M	580.30	562.27	553.98	555.21	553.98	551.01
SD	43.65	46.05	41.44	42.76	44.53	41.67
Below 25th percentile	16%	27%	33%	33%	31%	32%
Above 50th Percentile	61%	44%	36%	38%	36%	37%
Reading Comprehension						
M	589.75	578.32	574.64	571.01	569.32	569.57
SD	30.86	34.99	32.91	31.56	35.23	33.59
Below 25th percentile	12%	24%	29%	34%	34%	32%
Above 50th percentile	66%	51%	43%	39%	38%	36%

Note. There were no significant differences between the comparison groups as a function of the number of years the students has been in the school. First Grade: Word Reading ($F = 1.21$, $p \leq .19$); Reading Comprehension ($F = 1.98$, $p \leq .16$); Second Grade: Word Reading ($F = 3.02$, $p \leq .09$); Reading Comprehension ($F = .86$, $p \leq .43$). RITE = Rodeo Institute for Teacher Excellence.

Results (shown in Table 6) indicate that the number of years in the RITE program was significantly related to both Word Reading and Reading Comprehension. Follow-up contrasts for each grade-level analysis indicate that performance scores increased significantly as the number of years of experience in the RITE program increased. Therefore, all RITE students are performing at levels significantly higher than comparison students; and within the RITE program, students who finish first or second grade with more years of program experience outperform their program peers (see Tables 5 and 6).

Additional analyses were conducted to examine the effect of number of years in the same school on performance levels for comparison students. Because the number of years in the RITE program is confounded with the number of years the student remains in the same school, it was important to examine the effect of number of years

TABLE 6
First and Second Grade SAT9 Performance Predicted by Number of Program Years

Grade and SAT9 Subtest	*F*	*p*
First-grade Word Reading[a]	71.72	.0001
Follow-up contrasts		
Comparison versus 1 year	17.00	.0001
Comparison versus 2 years	143.42	.0001
1 year versus 2 years	69.90	
First-grade Reading Comprehension[a]	92.11	.0001
Follow-up contrasts		
Comparison versus 1 year	65.19	.0001
Comparison versus 2 years	168.67	.0001
1 year versus 2 years	40.58	.0001
Second-grade Word Reading[b]	27.65	.0001
Follow-up Contrasts		
Comparison versus 1 year	2.99	.080
Comparison versus 2 years	7.81	.0070
Comparison versus 3 years	63.10	.0001
1 year versus 2 years	14.78	.0001
1 year versus 3 years	75.00	.0001
2 years versus 3 years	27.25	.0001
Second-grade Reading Comprehension[b]	27.45	.0001
Follow-up contrasts		
Comparison versus 1 year	9.17	.0030
Comparison versus 2 years	21.97	.0001
Comparison versus 3 years	74.07	.0001
1 year versus 2 years	3.79	.0050
1 year versus 3 years	39.77	.0001
2 years versus 3 years	19.75	.0001

Note. Number of program years were used a predictor for this analyses. Comparison students were considered irrespective of the number of years in the same school based on analyses indicating no significant differences between these groups. SAT9 = Stanford Achievement Test–Ninth Edition.

[a]First-grade analyses, $df = 2, 6981$. [b]Second-grade analyses, $df = 3, 6{,}139$.

in the same school on performance within the comparison schools. Results indicated that there were no significant effects of number of years in the same school and first- and second-grade performance levels within the comparison schools (see Table 5). Based on this, all analyses were conducted collapsing the comparison students into one category (zero program years). These results strengthen the findings for the effects of number of years in the RITE program as they suggest that the effects are not simply an artifact of student stability within the same school environment.

As with the kindergarten analyses, first- and second-grade performance score distributions were also examined in terms of the percentages of students scoring below the 25th percentile and above the 50th percentile as a function of the number of years in the RITE program (see Table 5). As can be seen in this table, the greater the number of years in the program, the less likely students are to perform below the 25th percentile and the more likely they are to perform above the 50th percentile. The greatest performance differences are seen with students who at the end of second grade have had 3 years of program experience, or who at the end of first grade have had 2 years of program experience. Put another way, the program effects were greatest for students who began the program in kindergarten, and next largest for students who began the program in Grade 1.

Third-grade analyses. The final set of analyses of student outcomes compared third-grade TAAS performance of students in the RITE and comparison schools. (Third graders also took the SAT9, but evaluators did not have these scores available for analysis.) There are three groups of children who have participated in the RITE program who have completed the third grade. One group participated in the program for the entire 3 years (Grades K–2), the second participated in first and second grade (2 years), whereas the other participated for 1 year (second grade only). Hence, in the third-grade analyses, there are four groups of third-grade children being compared: those with 3 years in the program; those with 2 years, 1 year, and 0 years in the program; or comparison children.

The first set of analyses compared average performance levels on the reading portion of the TAAS across the four groups, and the second examined the percentage of children passing the reading portion of the TAAS, as shown in Table 7. Analyses indicated that average TAAS TLI scores for children in third grade who participated in the RITE program for 3 years (Grades K–2) were significantly higher than those for students who participated for 2 years, 1 year, or in a comparison school, as shown in Table 8. Furthermore, students participating for 2 years (first and second grade) have significantly higher average scores than those who participated in second grade only (1 year), who in turn have higher average scores than students from the comparison schools. As with the first- and second-grade analyses, an additional analysis was conducted to examine the effect of number of years in the same school on performance levels for comparison students. Results indicated that there was no significant effect of number of years in the same school on third grade TAAS reading TLI scores within the comparison schools (see Table 7). Based on this, all analyses were conducted collapsing the comparison students into one category (zero program years).

TAAS pass rates are calculated based on the students TLI score (70 or greater is equivalent to passing). Passing the TAAS at Grade 3 means a student has met the minimum expectations for the end of third grade. Analyses indicated a similar pat-

tern to that found with the average TLI score analysis mentioned earlier, $\chi^2 = 26.44$, $p < .001$. Specifically, as students spend more time in the RITE program, they are significantly more likely to meet or exceed the state-mandated reading skills requirement (see Table 7).

Longitudinal Analyses

The two groups of students who showed the highest performance levels in the Grade 1, Grade 2, and Grade 3 analyses mentioned earlier were those students who began the program in kindergarten. Although grade-level analyses inform us about the performance levels of students with varying years of experience in the RITE program, these analyses tell us little about the gains being made within a specific group of students over time. The longitudinal analyses examined the degree to which performance in a higher grade could be attributed to gains made the previous year, or whether there were additional gains being made above and beyond performance gains in previous program years.

Longitudinal analyses were conducted for the cohort of children who had participated in the entire span of the RITE program (Grades K–2) and for whom data were available in each year of the program (first kindergarten cohort beginning kindergarten in 1997–1998 and completing second grade in the 1999–2000 school year).

Analyses of performance from kindergarten through first grade were conducted by a series of models predicting SAT9 scores at the end of the first-grade year after controlling for performance levels on the Woodcock–Johnson Word Identification subtest in the kindergarten year. This analysis allowed us to examine differences in first-grade performance levels above and beyond the gains seen in the kindergarten year. Therefore, we can ascertain not only whether RITE children are performing at higher levels than comparison children at the end of first grade, but also whether ad-

TABLE 7
Third Grade Texas Assessment of Academic Skills Means and Percentage Passing by Number of Program Years

	RITE Years			*Comparison Years*		
Statistic	*3*	*2*	*1*	*3*	*2*	*1*
M	79.86	78.05	75.04	72.47	71.89	71.08
SD	16.44	18.26	21.22	22.66	24.02	23.46
Percentage passing	82%	79%	73%	68%	66%	65%

Note. There were no significant differences between the comparison groups as a function of the number of years the students have been in the school ($F = .24$, $p \leq .62$). RITE = Rodeo Institute for Teacher Excellence.

TABLE 8
Third-Grade Performance Predicted by Number of Program Years

TAAS	*F*	*p*
Reading TLI[a]	15.58	.0001
Follow-up contrasts		
Comparison versus 1 year	14.94	.0001
Comparison versus 2 years	27.11	.0001
Comparison versus 3 years	15.49	.0001
1 year versus 2 years	5.17	.0200
1 year versus 3 years	5.04	.0200
2 years versus 3 years	0.52	.4700

Note. Comparison students were considered irrespective of the number of years in the same school based on analyses indicating no significant differences between these groups. TAAS = Texas Assessment of Academic Skills; TLI = Texas Learning Index.

[a]Third-grade analyses, $df = 3, 5{,}040$.

ditional differential performance gains occurred in the first-grade year that were not attributable solely to the differential performance gains observed in kindergarten.

Analyses indicated that there were significant group differences (RITE vs. comparison) in students' SAT9 Word Reading and Reading Comprehension skill levels at the end of first grade after controlling for end of kindergarten Word Identification skills, as shown in Table 9. Therefore, not only do RITE first-grade students with 2 years of program experience end first grade at higher performance levels, these students show differential performance gains across the first-grade year that cannot be attributed solely to gains seen in the kindergarten year.

Analyses of performance from first through second grade were also conducted for this cohort by a series of models predicting SAT9 scores at the end of the second-grade year after controlling for performance levels on the same SAT9 subtests at the end of the first-grade year. These analyses examined differences in performance at the end of second grade that were above and beyond any gains seen in the first-grade year. Therefore, these analyses examined whether RITE children performed at higher levels than comparison children at the end of second grade, and whether additional development occurred in the second-grade year that was not attributable solely to performance gains in the first grade.

Results indicated that there were no significant differences in RITE and comparison students' average Word Reading or Reading Comprehension scores at the end of the second-grade year after controlling for end of first-grade performance (see Table 9). Although all students' Word Reading and Reading Comprehension performance levels increased over the second-grade year, results suggest that performance at the end of second grade for the cohort with 3 years in the RITE pro-

gram has more to do with gains made in previous years rather than gains made in the second-grade year.

Teacher Implementation Analyses

Teacher correction techniques. In observing teachers' corrections of errors in students' responses, the first piece of information recorded was whether an error was made. If so, the observer then recorded whether the teacher provided a full, partial, or no-correction for the erroneous response. The percentage of errors was calculated by dividing the number of errors observed by the number of responses observed. On average, errors were observed in 15% of the student's responses in the fall and in 10% of the responses in the spring. At each time point, the percentage of errors observed across all classrooms ranged from 1% to 30%.

The number of years a teacher is in the RITE program will likely influence their skill in implementing the key aspects of the program. To examine this, teacher corrections were examined as a function of the number of years of experience that the teacher had with the RITE program, as shown in Table 10. In the beginning of the year, fourth-year teachers showed lower use of full corrections than all other teachers. However, by the end of the year, there were no differences in full correction usage based on teaching experience.

Examination of the teachers' responses to students' responses showed little to no variation in the percentage of time the teacher provided praise or verification over the course of the year. Overall, RITE teachers praised student responses an average of 20% of the time, and provided verification of responses an average of 30% of the

TABLE 9
End of Grade-1 Performance Controlling for End of Kindergarten Performance in K–2 Cohort

SAT9 Subtest	*Predictor*	*F*	*p*
End of Grade 1[a]			
Word Reading	WJ Word Identification end of Kindergarten	11.83	.0007
	Group (RITE vs. COMP)	17.42	.0001
Reading Comprehension	WJ Word Identification end of Kindergarten	7.01	.0090
	Group (RITE vs. COMP)	22.09	.0001
End of Grade 2[b]			
Word Reading	Word Reading end of Grade 1	187.40	.0001
	Group (RITE vs. COMP)	.80	.4400
Reading Comprehension	Reading Comprehension end of Grade 1	299.60	.0001
	Group (RITE vs. COMP)	1.32	.3300

Note. COMP = comparison; SAT9 = Stanford Achievement Test–Ninth Edition; WJ = Woodcock–Johnson.

TABLE 10
Percentage of Teacher Corrections Over Time and by Number of Program Teaching Years

	Time Point	
Program Years	*Fall*	*Spring*
1 year	60%	76%
2 years	57%	77%
3 years	67%	76%
4 years	46%	77%

time. Furthermore, there were no differences in the average use of praise or verification as a function of the number of years teaching in the RITE program.

These results are not surprising given that the increased use of programmatic correction techniques was a primary focus of the RITE training program in the current year. Based on the evaluation of teacher behaviors in prior years, RITE trainers this year increased efforts to improve teachers' use of full corrections in response to student errors. The average gains seen across the year for all teachers were consistent with this general emphasis. At the same time, emphasis on praise and verification were reduced as previous years' evaluations have not found these two teacher behaviors to be as strongly linked to student outcomes. Not surprisingly, these teacher behaviors are relatively stable over the current year.

Trainer Intervention Analyses

General classroom teaching skills. Table 11 shows the average pre and postintervention ratings for general classroom teaching skills, as well as the average level of reported intervention by number of years teaching in the RITE program. First- and second-year RITE program teachers' general teaching skills were rated as more problematic than third- and fourth-year teachers. On average, trainers reported providing more intervention to these teachers than to either third- or fourth-year teachers. Furthermore, first- and second-year teachers were rated by trainers as showing the most improvement in these teaching skills over the course of the year. Although trainers reported providing more intervention with teachers where ratings were more problematic, it was encouraging to note that the correlation of reported intervention levels with the observed decrease in these behaviors was negative—meaning that for all teachers, the more intervention provided, the more problematic behaviors decreased in the areas of classroom management, organization, and disciplinary technique or behavior management (–.59, $p \leq .0001$). Although en-

TABLE 11
Trainer Reported Problems and Intervention Levels by Number of Program Teaching Years

	Number of Years of RITE Teaching			
Type of Rating	*1*	*2*	*3*	*4*
Classroom management				
Pre-intervention problems	2.0	1.8	1.2	1.3
Post-intervention problems	1.4	1.3	1.2	1.3
Level of intervention	2.4	2.2	1.6	1.8
Programmatic teaching				
Pre-intervention problems	1.9	1.8	1.4	1.3
Post-intervention problems	1.2	1.2	1.2	1.1
Level of Intervention	2.8	2.6	1.8	1.8

Note. RITE = Rodeo Institute for Teacher Excellence.

couraging, because the same person made both ratings (intervention and observed improvement), it is also possible that the correlation simply reflects rater bias. To examine this possibility, the relation between reported level of general classroom teaching skills and examiner observation of behavioral interruptions in the classroom was examined. The correlation between the two was significant, indicating that the more reported intervention on the part of the trainer, the greater the decrease in the behavioral interruptions observed by examiners in the classroom over the course of the year (–.29, $p \leq .01$). Here, the ratings of teacher behaviors and level of intervention required were made by trainers, whereas the observations of teacher behaviors used to assess teacher behavior change were made by classroom observers who work for the evaluation team. These individuals are not involved in the rating or training of teachers and do not have contact with the trainers. In that sense, the observations used to measure teacher behavior and teacher behavior change are made independently of the trainers' ratings of teachers. Therefore, it seems that trainer intervention with general classroom management skills is related to the teacher's ability to better manage the behavior of the children in the classroom.

Program-specific teaching skills. Table 11 also shows the average pre and postintervention ratings for program-specific teaching skills, as well as the average level of reported intervention. On average, first- and second-year RITE program teachers' program-specific skills were rated as more problematic than the skills of third- and fourth-year teachers, and trainers reported significant improvement in these skills after intervention for first- and second-year teachers. In addition, by the end of the year, trainers were reporting, on average, relatively few problems with these skills for all teachers. Although trainers reported providing more intervention with teachers where ratings were more problematic, it was encouraging to note that the correlation of reported intervention levels with the reported decreases in problematic program im-

plementation skills was significant—meaning that for all teachers, the more intervention provided, the greater the rated decrease in problematic behaviors in the areas of program implementation skills (–.54, $p \leq .0001$). Again, although this relation was encouraging, the fact that both of these ratings (intervention and observed improvement) were made by the trainer leaves open the possibility that the correlation is an artifact of the rater. To examine this possibility, the level of intervention required to improve program implementation skills as rated by the trainer was correlated with teacher correction techniques in the classroom as rated by the classroom observer. The correlation between the two was significant and positive, indicating that the more reported intervention on the part of the trainer, the greater the increase in the teachers' use of full corrections over the course of the year (.21, $p \leq .01$). Therefore, it seems that trainer intervention with program implementation skills is related to teachers' increased use of full correction techniques in the classroom.

It is interesting to note that this pattern did not hold true for teachers' implementation of verification responses. There was no relation between reported levels of intervention for program-specific skills and the degree to which teachers used verification in their responses to children. It may be the case that when reporting levels of intervention for program implementation techniques, trainers focused more heavily on teacher correction behaviors than teacher responses to students' correct responses (or praise and verification responses).

LINKING TEACHER IMPLEMENTATION TO STUDENT PERFORMANCE

An important element of any program's success lies in the degree to which implementation of the key components of the program relates to desired outcomes for students. To gain support for the specific program being used, it is important to first establish that key components of the program are indeed related to desired outcomes, and that the degree to which implementation of the key components is followed is correlated with higher desired outcomes. This type of evidence provides strong support for the specific program as a route for obtaining desired outcomes.

In this evaluation, the degree to which full implementation of program-specific components were related to increased student achievement was examined. This portion of the evaluation focused on teachers usage of full correction techniques at the beginning and end of the school year. Simultaneous examination of the relation of usage at these two time points to student outcomes allows for the determination of the relative influence of implementation levels at the beginning and end of the year.

The relations between fall and spring levels of teacher corrections were examined simultaneously for each student outcome in each grade. In all models, the interaction between fall and spring levels of teacher corrections was included to allow for the possibility of different outcomes based on the difference in the level of corrections across the year. Results indicated that teacher corrections related

significantly to children's performance levels, and that the pattern of the relation was similar across all grades.

Results indicated that teachers' use of correction techniques in both the fall and spring was related to students' performance on the majority of skills, as shown in Table 12. Specifically, the higher the level of usage, the higher the students' performance levels. The interaction between fall and spring corrections was also significant in all models. Follow-up analyses indicated that students of teachers who used low levels of correction over the course of the year performed at significantly lower levels than all other students. Therefore, teachers' high use of correction techniques for all or at least part of the school year was more effective than no use of full correction techniques.

DISCUSSION

Student Outcomes

This evaluation indicated that the RITE program was very successful at increasing the reading abilities of students in at-risk schools and who would likely themselves be at risk for reading difficulties. Children who began the RITE program early and

TABLE 12
Student Performance Predicted From Fall and Spring Teacher Corrections

Performance Measure	*Predictor*	*F*	*p*
Kindergarten[a]			
Word Reading	Fall corrections	6.13	.01
	Spring corrections	5.68	.02
	Interaction	6.01	.01
Grade 1[b]			
Word Reading	Fall corrections	15.70	.0001
	Spring corrections	20.60	.0001
	Interaction	28.04	.0001
Reading Comprehension	Fall corrections	28.44	.0001
	Spring corrections	20.03	
	Interaction	35.28	
Grade 2[c]			
Word Reading	Fall corrections	14.20	.0001
	Spring corrections	19.80	.0001
	Interaction	16.76	.0001
Reading Comprehension	Fall corrections	25.89	.0001
	Spring corrections	19.68	.0001
	Interaction	27.54	.0001

[a]Kindergarten analyses, $df = 3, 1{,}459$. [b]First-grade analyses, $df = 3, 1{,}646$. [c]Second-grade analyses, $df = 3, 1{,}877$.

who spent more years in the program outperformed their schoolmates with less program experience, those who began the program later, and those who never participated in the program (comparison school students). The most profound effects of the RITE program were seen in the first 2 years of schooling, especially when students began the program in kindergarten. By the end of kindergarten, students showed prereading skill development levels greater then their nonprogram peers; they also demonstrated greater gains in these skills over the course of the kindergarten year. At the end of first grade, children with 2 years in the program again outperformed their peers, both those with less program experience and those who had not participated in the program. Furthermore, these first graders also showed differential gains during the first-grade year that could not be accounted for by the gains experienced in kindergarten alone.

In second grade, RITE students with previous experience in the program continued to perform at higher levels than their peers with less program experience and comparison students. However, these children did not show differential gains across the second-grade school year. Therefore, in second grade, growth rates in reading skills were comparable for RITE and comparison school children, whereas overall performance level differences between the groups were maintained.

Based on these findings, we conclude that the program has accelerated students' development of prereading and reading skills. By second grade, the acceleration of this development has slowed, such that skill development in second grade continues at rates that are comparable to those of nonprogram students. Therefore, we also conclude that the second-grade program as currently implemented does not fully capitalize on the performance gains experienced by children who participate in the program in kindergarten and first grade. However, it is important to recognize that students in the comparison schools have not caught up to the RITE program children by the end of second grade. In fact, third-grade students who have participated in the RITE program were significantly more likely to pass the minimum skills requirement on the reading section of the TAAS than were students in the comparison schools. Nevertheless, RITE must consider steps that can be taken to further improve outcomes for students in second grade and beyond, including finding ways to strengthen the impact for students whose first year in the program is in second grade, and ways to better capitalize on the gains made in kindergarten and Grade 1. Currently, the leadership of the RITE program is considering enhancements to the language and literacy components of the program, and in particular, working with teachers to increase the amount of book reading and language development activities employed in Grades K–2.

Teacher Implementation and Trainer Support

At the beginning of the year, first-year teachers were rated as having more problematic general classroom and program-specific teaching skills. Not surprisingly,

trainers on average also reported providing first-year teachers with more support (intervention) with general classroom teaching as well as with program-specific teaching skills. Analyses within this year's evaluation indicated that the more intervention the trainer reported providing to a teacher, the more improvement there was in the observed teaching skills discussed earlier (behavior management and teacher corrections). Most notable was the success of the RITE trainers focus on full correction techniques and the gains seen in the majority of the teachers' implementation of these techniques over the course of the school year. Because the level of intervention was related to observed positive development of these skills, and, as we saw earlier, there was still room for the development of these skills, it would be important to continue high levels of intervention with all teachers, regardless of the number of years in the program.

Student Performance and Teacher Implementation

Implementation of the more advanced teaching techniques required by the RITE program was significantly related to student performance. Teachers who showed higher levels of implementation all year or part of the year had students who were performing at significantly higher skill levels than teachers who showed low levels of implementation all year.

These findings are important in that they indicate that teaching techniques that are specific to delivery of the RITE program are related to better student performance. Furthermore, taken with the previous discussion of the effects of trainer intervention, these results also support the importance of the training component of the program. Specifically, we saw that trainer intervention was related to teachers' improved adherence to program teaching techniques, and that teacher adherence to program teaching techniques resulted in better student outcomes. These results close the trainer–teacher–student feedback loop by showing that teacher behavior relates to student performance. Insofar as room remains for improvement in teacher adherence to program teaching techniques, trainers' support of teachers must be continued and strengthened. As trainers increase their support of teachers, the RITE program can expect more improvement in teachers' adherence to program teaching techniques, and as a result, greater gains in student performance can be expected.

ACKNOWLEDGMENTS

The RITE program and the research associated with it are supported by The Houston Livestock Show and Rodeo.

We wish to thank the RITE teachers, administrators, and trainers; as well as the individual students who take part in the RITE program.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Brady, S., & Moats, L. C. (1997). *Informed instruction for reading success: Foundations for teacher preparation.* Baltimore: International Dyslexia Association.

Englemann, S., & Bruner, E. C. (1995). *The SRA reading mastery rainbow*. New York: McGraw-Hill.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 85,* 112–126.

Juel, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly, 31,* 268–288.

Leslie, L., & Allen, L. (1999). Factors that predict success in an early literacy intervention project. *Reading Research Quarterly, 34,* 404–424.

Lipson, M. Y., & Wixson, K. K. (1997). *Assessment and instruction of reading and writing disability: An interactive approach* (2nd ed.). New York: Plenum.

Little, R. J., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: Wiley.

National Reading Panel Report. (2000, April 13). Testimony of Duane Alexander, Director, National Institute of Child Health and Human Development before the Labor, Health and Human Services, and Education Subcommittee, Senate Appropriations Committee, Washington, DC.

Pflaum, S. W., Walberg, H. J., Karegaines, M. L., & Rahsher, S. P. (1980). Reading instruction: A quantitative analysis. *Educational Researcher, 9*(7), 12–18.

Shafer, J. L. (1997). *Analysis of incomplete multivariate data.* London: Chapman & Hall.

Snider, V. E., & Tarver, S. G. (1987). The effect of early reading failure on acquisition of knowledge among students with learning disabilities. *Journal of Learning Disabilities, 20,* 351–356.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Stahl, S. A., McKenna, M. C., & Pagnucco, J. R. (1994). The effects of whole language instruction: An update and reappraisal. *Educational Psychologist, 29,* 175–186.

Texas Education Agency (TEA). (1990). *Texas assessment of academic skills: Technical report.* Austin: Author.

Texas Education Agency (TEA). (1998). *Texas primary reading inventory: Technical manual.* Austin: Author.

Wharton-McDonald, R., Pressley, M., & Hampston, J. (1998). Literacy instruction in nine first-grade classrooms: Teacher characteristics and student achievement. *The Elementary School Journal, 99*(2), 101–128.

Woodcock, R. W., & Johnson, M. B. (1979). *Woodcock–Johnson Psychoeducational Battery–Revised.* Allen, TX: DLM.