A School District Comparison of Reading Achievement Based on Three Reading Programs

Sharon McCollum, Mary Nell McNeese, Ronald Styron, and David E. Lee

Abstract: This study compared the reading achievement levels of 323 third grade students from a Caribbean school district receiving instruction from three different programs. Students were identified as at risk with a 95% minority enrollment, 100% free lunch and transportation, and the lowest NAEP test scores in the nation. Total standardized test scores from the Wide Range Achievement Test (WRAT)-Expanded Edition were compared according to instructional methods. The results of an ANOVA showed significant differences. Statistically, the Success for All Group achieved the highest mean score, while there were no significant differences in the mean scores between the Direct Instruction Group and the Basal Reader Group. The importance of this study lies in its effort to analyze available data on three modes of reading instruction. The school district should select one reading program and institute it districtwide after longitudinal quantitative and qualitative data are collected. This study is only the first step in that direction.

The many factors that place young children at risk educationally include poverty, language barriers, learning disabilities, minority ethnic group membership, or a combination of such factors (Baas, 1991). As a result, “the challenge for educational researchers and practitioners is to identify practical and effective means of reducing such students’ chances of academic failure, grade retention, and dropping out of school at an early age” (Ross, Smith, Slavin, & Madden, 1997, p. 171).

As Englemann (1999) reported, school failure for at-risk students results largely from the fact that all children are expected to learn a specified battery of skills in a specified number of years. This comparison may be unfair for at-risk children because they may take longer to master those skills. They enter first grade substantially behind in reading, language, and number skills (p. 77).

Failure to develop reading skills during the primary school years intensifies as the student progresses through the grades. Juel (1988) reported that approximately 88% of first-grade students whose performance scores were in the lowest quartile in reading comprehension remained at performance levels below the 50th percentile through the fourth grade. Similar findings reported by others indicate that “students who have been poor readers in the early elementary years remain poor readers throughout school” (Carlson & Francis, 2002, p. 142).

The Comprehensive School Reform Movement (CSRM) promotes the idea that student achievement occurs most frequently when there is an intensive effort to make positive, academically-focused, schoolwide changes. Those “students in schools working with whole-school reform tended to achieve greater gains than students in schools attempting various pull-out programs” (Wested, 2005, p. 5). However, in spite of the promise shown by these programs, educators continue to be puzzled by the large number of children with severe problems in reading.

The purpose of the study reported here was to compare the reading standardized achievement test scores of third-grade students who received reading instruction using three different programs: Success for All (SFA), Direct Instruction (DI), and Basal Reader (BR). The backgrounds, advantages, and criticisms of each reading program were also considered.

Success for All

SFA, a school restructuring program developed by Robert Slavin and Nancy Madden (2000) of Johns Hopkins University, was designed to address the needs of at-risk students in high poverty schools. Key features of the program are

(a) a research-based instructional component focused on the development of literacy and oral language skills, (b) cooperative learning, (c) assessment of educational progress every eight weeks, (d) homogeneous ability grouping for reading instruction, (e) one-on-one tutoring, (f) a family support team, and (g) a full-time program facilitator to administer the program. (Urdegar, 2000, p. 1)

In contrast to traditional strategies, which often emphasized pulling disadvantaged students out of regular classes to receive limited tutoring, SFA was designed as a comprehensive program grounded on two essential principles: prevention and immediate, intensive intervention (Slavin et al., 1996; Slavin, Madden, Karweit, Dolan, & Wasik, 1992). SFA proposes that at-risk students are more successful when their academic deficiencies are addressed.
early. A reform model, it is highly specified and comprehensive with respect to implementation guidelines and material for students and teachers. Almost all materials for students are provided, and teachers are expected to follow SFA lesson plans closely (Madden, Livingston, & Cummings, 1998).

The claims about the effectiveness of SFA went unchallenged until Walberg and Greenberg (1999) argued that independent evaluations by Venezky (1997) and Jones, Gottfredson, and Gottfredson (1997) showed that SFA produced gains only in kindergarten and first grade. These researchers concluded that at these early grade levels, it is easy to produce gains using a wide variety of much cheaper, simpler methods.

**Direct Instruction**

According to the American Federation of Teachers (2003), the oldest version of Direct Instruction, DISTAr, was developed in the 1960s as part of Project Follow Through, a component of President Johnson’s War on Poverty. DISTAr achieved some level of success; however, it was heavily criticized for being too rigid, for concentrating too heavily on the basics, and for poor implementation practices (Adams & Englemann, 1996). The original DISTAr program, which has been expanded and enriched, is what is now termed DI.

This model “is a comprehensive system of instruction that integrates effective teaching practices with sophisticated curriculum design, classroom organization and management, and careful monitoring of student progress, as well as extensive staff development” (Stein, Carnine, & Dixon, 1998, p. 227). According to the American Federation of Teachers (2003), it “is a highly structured instructional approach, designed to accelerate the learning of at-risk students. Curriculum materials and instructional sequences attempt to move students to mastery at the fastest possible pace” (p. 1). Specifically, this approach integrates effective teaching practice such as monitoring student performance, providing corrective feedback, increasing academic engaged time through the use of small group instruction, and unison responding. The effective teaching techniques must be tied to well-designed, generalizable instructional strategies in order for students to succeed academically (Stein et al., 1998, p. 228).

Adams and Englemann (1996) have identified 54 well-designed studies that compared DI models to other instructional approaches. Results showed that 87% of the post-treatment means favored the DI model, compared to only 12% favoring non-DI approaches. Sixty-four percent of the statistically significant outcomes favored the DI model, compared to only 1% of the outcomes favoring nondirect approaches, and 35% showed no difference among the approaches.

**Basal Reader**

Adopted BR series have been used as one component of elementary school language arts curriculums in the United States for decades. BRs popularized the “look-say" method of reading instruction in the 1950s. The most popular of the early BRs utilizing this method was Scott Foresman’s “Sally, Dick and Jane.” The focus of BRs was repeated practice with the same small set of vocabulary.

Criticism of BR programs focused on the lack of attention to systematic phonics instruction (Hoffman, Sailors, & Patterson, 2004), a problem addressed in the 1970s and 1980s; however, critics then contended that teachers became overly reliant upon skill-oriented workbooks and manuals and that their students still performed lower on national reading achievement assessments. Other criticism included that BRs failed to provide purposeful reading and underrepresented minorities and existing racial conflict in the stories (Pirofski, 2003).

Despite problems associated with basal reading programs, BRs have been shown to be helpful in developing reading proficiency in most children. Current BR programs consist of a full complement of materials, including comprehensive teacher guides; practice workbooks; testing materials; and instructional aids such as charts, word cards, “Big Books,” game boxes, supplementary paperback library books, dictionaries, and reproducible masters for classroom handouts. Dechant (1991) reported that 95% or more of elementary school teachers use a BR approach, even though that percentage is declining.

**Method**

**Participants**

The study of these three methods reported here was conducted in a public school district in one of the U.S. Territories in the Caribbean Sea, which includes three islands and numerous keys. The district had an enrollment of 18,700 students, with an annual per-pupil expenditure of $6,478. All of its schools were classified as Title I schools, with their students receiving both free lunch and transportation. The school district was characterized by low-achieving students. In 1992 the average National Assessment of Educational Progress (NAEP) reading score for the district was 171 compared to a national average score of 215, and in 2002 the average NAEP reading score was 179 compared to a national average of 217. The schools serviced children, many of whom came from single parent households, non-English speaking households, households speaking an indigenous dialect, and households where the students were the first in their families to attend an American school. The ethnic background of the students was Afro-Caribbean.

Although there were 12 elementary schools in this district, this study was delimited to seven based on their choice of the three reading programs under investigation. Four hundred Wide Range Achievement Test (WRAT)-Expanded 2003-04 reading standardized test results were obtained from the participating schools in accordance with the district’s policy and procedures. A total of 323 test results were useable, 77 being discarded because a different level of the WRAT-Expanded Group Assessment had been administered to those students.

**Materials**

The WRAT-Expanded Assessment, chosen because it is the only standardized instrument that had been administered to the public school children in the district in the last 5 years, measures those reading abilities important for understanding printed material beginning with reading words and sentences,
then continuing with the comprehension of various types of reading passages of ever-increasing complexity appropriate for proficient readers at the elementary and secondary levels. (Robertson, 2001, p. 5)

Test items assess three aspects of comprehension: Literal, Inferential, and Word Meaning in Context.

The test’s psychometric properties of reliability, error, and validity have been assessed. Robertson (2001) reported that the test-retest reliability of WRAT-Expanded, as measured by the Kuder-Richardson Formula 20 (KR-20), was .89, which is acceptably reliable. As he explains, “All tests contain error to some degree. The Standard Error of Measurement (SEM) for the WRAT-Expanded tests is 5.1” (Robertson, 2001, p. 31). To gain construct and content validity, the tests were research literature-based, and test experts made suggestions for revisions, which were subsequently incorporated into them. Robertson concluded that the WRAT-Expanded norms were generally consistent with those of other tests normed at different times on different samples of individuals, saying that “these findings suggest that WRAT-Expanded users can have confidence that the norms represent a sample of examinees generally similar to the norming samples of other widely used achievement and cognitive measures” (Robertson, 2001, p. 52).

Procedures

All students received instruction in one of the three respective programs for 4 years spanning grades K-3. Eighty-seven, or 26.9%, of the students received instruction via SFA; 126, or 39%, received instruction via DI; and 110, or 34.1%, received instruction via BR.

The Reading Roots and Wings program was used for SFA instruction. Students were assessed and regrouped according to their reading level each school quarter when their teachers, administrators, and SFA coaches met to review their progress. Interventions were implemented as needed. Ongoing coaching and support were available to the schools through telephone meetings and site visits. The SFA program offered instruction on listening comprehension, teamwork (Treasure Hunts), writing (Adventures in Writing), editing (Two-Minute Edit), and a book club.

The SRA/McGraw-Hill program was used for DI instruction. The implementation of DI entailed language instruction using the Language for Learning program, reading instruction using the Reading Mastery and/or Corrective Reading programs, spelling instruction using the Spelling Mastery program, and writing instruction using the Reasoning and Writing program. No coaching or instructional supports were provided to schools that implemented this program of instruction.

The district used Literature Works: An Integrated Approach to Reading and Language Arts for BR instruction. Literature Works for grades K-6 was designed to motivate students through a wide range of reading materials. The anthologies in Literature Works, entitled Collections, centered on themes directed towards student interests. These themes were presented in both fiction and nonfiction works. A Theme Launch was provided for all grade levels, providing opportunities for students to preview the theme, develop a common language, build background, and set learning goals. No coaching or instructional supports were provided to schools that implemented this program of instruction.

The WRAT-Expanded Reading Assessment was administered to the third-grade students in March 2003 by the classroom teachers under the supervision of the school counselor, school administrators, and central office personnel. After securing district approval, the school counselors reported the total reading test standardized scores to the researchers. The data were entered in the Statistical Program for the Social Sciences (SPSS), version 11.

Results

The WRAT-Expanded Reading Achievement test mean for the entire sample was 92.2. The mean for the SFA students was 94.8, the highest of the three reading programs, while DI students had a mean of 90.2, the lowest of the programs compared, and BR students had a mean of 92.3. Both SFA and the BR had means above the sample. The sample standard deviation was 11.99. Figure 1 below compares the WRAT-Expanded standardized reading achievement score means for the three groups of third grade students.

Table 1 presents the descriptive statistics for the standardized reading achievement scores among the three groups in the study.

A one-way analysis of variance (ANOVA) was conducted to evaluate the differences among students from the three reading programs on the total WRAT-Expanded reading test standardized scores. The independent variable, the reading program factor, included the three programs, the SFA, the DI, and the BR. The dependent variable was the total WRAT-Expanded reading test standardized achievement scores. The ANOVA was significant, $F(2, 320) = 3.85, p = .02$. The strength of the relationship between the reading program factor and the total WRAT-Expanded reading test standardized scores as assessed by $\eta^2$ was small, with the reading factor accounting for 2% of the variance of the dependent variable. Even though this effect was small, it did show practical significance based on Cohen’s rule of thumb (Kirk, 1995) and was either higher or comparable to the studies of the three programs individually.
Tukey post-hoc comparisons were conducted to evaluate pairwise differences among the means since the Levene's test for equality showed that equal variances could be assumed. There were significant differences between the SFA group and both the DI and the BR groups, with the SFA group showing the highest mean scores. There were no significant differences, however, between the DI and BR groups, suggesting that the DI and the BR groups were statistically equal. The results of the data analysis showed that students in the SFA reading program scored significantly higher than students in the DI or BR reading programs, although students in the latter two programs did not score significantly different from each other.

Discussion

The data indicated that there was a statistically significant difference in standardized reading achievement scores among the SFA, the DI, and the BR groups. The SFA group had the highest standard reading achievement score. Students who received 4 years of instruction in the SFA reading program attained significantly higher standard mean scores than students who received reading instruction for a similar period of time with DI or BR. The reading programs will be discussed individually.

Success for All

SFA was designed as a comprehensive program grounded on two essential principles: prevention and immediate, intensive intervention. In SFA, attention is focused on providing every student the support needed to be a successful reader by the end of the third grade. The program recognizes the different ways and rates that students learn to read. Therefore, it offers a variety of support systems (Slavin et al., 1992).

The creators of the SFA program indicated that the program's success was dependent upon those who implemented it (Hill, 1998). Decidedly, the schools contributed to the positive or negative impact of all of the reading programs cited in this study, with teacher acceptance, teacher morale, teacher and staff training, and administrative support being but a few of the variables that contributed to the level of accomplishment in each. In addition, the teachers were provided added supports to help with the successful implementation of all three programs. The teachers in this school district appeared to support and be committed to the implementation of these programs, including SFA.

Direct Instruction

Proponents of the DI model contend that it is “a comprehensive system of instruction that integrates effective teaching practices with sophisticated curriculum design, classroom organization and management, and careful monitoring of student progress, as well as extensive staff development” (Stein et al., 1998, p. 227). The American Federation of Teachers (2003) evaluated DI as “a highly structured instructional approach, designed to accelerate the learning of at-risk students. Curriculum materials and instructional sequences attempted to move students to mastery at the fastest possible pace” (p. 1). Adams and Engelmann’s (1996) metanalysis of 34 studies found DI to be the most effective instructional reading program.

Contrary to these research literature findings, the study reported here found students who received DI reading instruction to be on par with their BR cohorts, while performing significantly behind their SFA counterparts. DI was designed for substantially the same target population of challenged readers as those for whom SFA was created. Becker (2001) and Engleman (1999) contended that DI is a valuable intervention when teaching reading to disadvantaged students. Under circumstances different from those in this study, the DI intervention may have produced positive reading gains similar to the SFA intervention.

Basal Reader

The BR was designed to increase reading ability and facilitate language arts skills in young readers by introducing children to selected series readings which gradually become more difficult. Traditionally, BR reading instruction has been the predominant method of reading instruction. However, Hoffman, Sailors, and Patterson (2004) found that BRs were not the most effective with minority students, such as the target population for this study. This study did not contradict those findings.

Delimitations and Limitations of the Study

The study was subject to the following delimitations:

1. The study was delimited to one of the U.S. Territories in the Caribbean Sea. This location was selected as an excellent example of a low-performing minority school district.
2. Additionally, the study was delimited to third grade students who received instruction through the SFA, DI, or BR reading programs for 4 years. This length of instruction was chosen based on the uniform assessment of reading achievement levels by the standardized WRAT-Expanded Test at the end of that period.

The following limitations restrict this study:

1. The data were collected by the school district, and, therefore, the researchers were limited to the data made available through the Superintendent’s Office. The researchers did not directly participate in the data collection.
2. Since the WRAT-Expanded Test was administered to students at each individual school site, the researchers were limited by the testing conditions selected by each school.
3. The students were selected based on their 4-year participation in the reading program at one of the seven schools in the sample.
Neither the students nor the schools in the study were randomly assigned to treatment groups. The absence of random selection reduces the meaning and the generalizability of this study; however, it was not feasible to randomly assign these students or schools to specific reading program treatment groups.

4. The study did not measure actual gains in reading ability from pretest to posttest. Since the students were not tested prior to the reading program, there may have been preexisting differences among students and/or schools that unfairly biased the WRAF-Expanded Test results in favor of the SFA reading program over the DI and the BR programs.

5. The implementation of the three reading programs was not standardized, and the teaching interventions were not assessed for validity or reliability. It is possible that the SFA schools had more gifted teachers and/or administrators than their cohorts. No measures were conducted to ensure that the instruction for any of the three interventions was actually implemented as designed.

6. The schools implementing the SFA intervention were provided program coaches to assist in the successful implementations of that program. These schools might have received additional training not afforded to those schools implementing the alternative reading interventions.

Implications for Educational Leaders
According to Ediger (2002),

the first R (reading, writing, and arithmetic) is vital for pupils to develop knowledge and skill since reading cuts across the curriculum and is highly important in society. Thus, reading is used in each and every academic area of the school curriculum. Teachers need to excel in reading instruction. (p. 1)

To aid in determining the best type of instruction, Guthrie, Schaefer, Von Secker, and Alban (2004) advocated the need for studies on characteristics of school reading programs producing reliable increases in student achievement. At the same time, however, they acknowledged the challenges of “detecting effects on achievement of reading programs when the school is the unit of analysis” (p. 2).

School reading programs are increasingly under scrutiny by school-based administrators, central office personnel, and policymakers. In the 1990s, voters expressed their dissatisfaction with low student test scores and the persistent achievement gap between Caucasians and most minority students. As a result, the No Child Left Behind Act of 2001 (NCLB) was passed. This legislation requires states to develop annual assessments aligned to state standards and to use achievement on these assessments as the primary measure of district and school accountability. NCLB is intended to ensure that all schools make adequate yearly progress (AYP) toward having all students proficient in reading and mathematics by 2014. The law also requires states to have in place a statewide accountability system that applies to all public schools, including charter schools (Learning First Alliance, 2004).

Schools are required to demonstrate that students are making adequate yearly progress, and failure to show AYP has dire consequences, including releasing that failure to the public, giving families the option of transferring their children to other schools, losing federal funding, instituting new curricula, replacing district personnel, appointing a trustee to run the district, and/or district restructuring.

As a result of NCLB, today, more than ever, it is essential that teachers, administrators, and district policymakers understand the characteristics of an effective schoolwide reading program. It has been the intent of this research to provide data to those teachers, administrators, and policymakers on three different reading programs in a low-performing minority school district.

Suggested Future Research
It is recommended that in the future, this research should be replicated over a longer period of time (such as 5 years) to evaluate the reading programs longitudinally. The authors of this work also recommend additional large-scale controlled studies to measure the efficacy of these programs. These studies should examine the effects of the limitations and delimitations on the measurement of the reading gains associated with each program, including the validity/reliability of those interventions and their standard implementation. Finally, there is a need for a pre-post comparison of all three reading programs.

References


---

**Authors**

**Sharon A. McCollum**, Ph.D., is a high school principal on a Caribbean island. Sharon’s research interests include: Reading Achievement, School Improvement, and Leadership. saminstvi@hotmail.com

**Mary Nell McNeese**, Ph.D., is an assistant professor of Educational Statistics and Foundations in the Department of Educational Leadership and Research at the University of Southern Mississippi. Mary Nell’s research interests include: National Center of Educational Statistics (NCES) Large scale Longitudinal Databases, Technology Integration, and Adult Learners. mary.mcneese@usm.edu

**Ronald A. Styron**, Ed.D., is an associate professor of Educational Administration and the chair of the Department of Educational Leadership and Research at the University of Southern Mississippi. Ron’s research interests include: Leadership, Principal Preparation, School Improvement, and Administrative Applications of Technology. ronald.styron@usm.edu

**David E. Lee**, Ed.D., is an assistant professor of Educational Administration in the Department of Educational Leadership and Research at the University of Southern Mississippi. David’s research interests include: Motivation, Communication, K-12 Administration, and School Improvement. david.e.lee@usm.edu