

Effects of Success For All on TAAS Reading Scores

A Texas Statewide Evaluation

The statewide data for Texas reported here show that Success for All schools are significantly and substantially closing the gap in TAAS reading performance between themselves and the far less impoverished schools in the rest of the state. The authors found this to be particularly true for African American and Hispanic students, for whom the gap with white students closed significantly more than it did for African American and Hispanic students in other schools throughout the state.

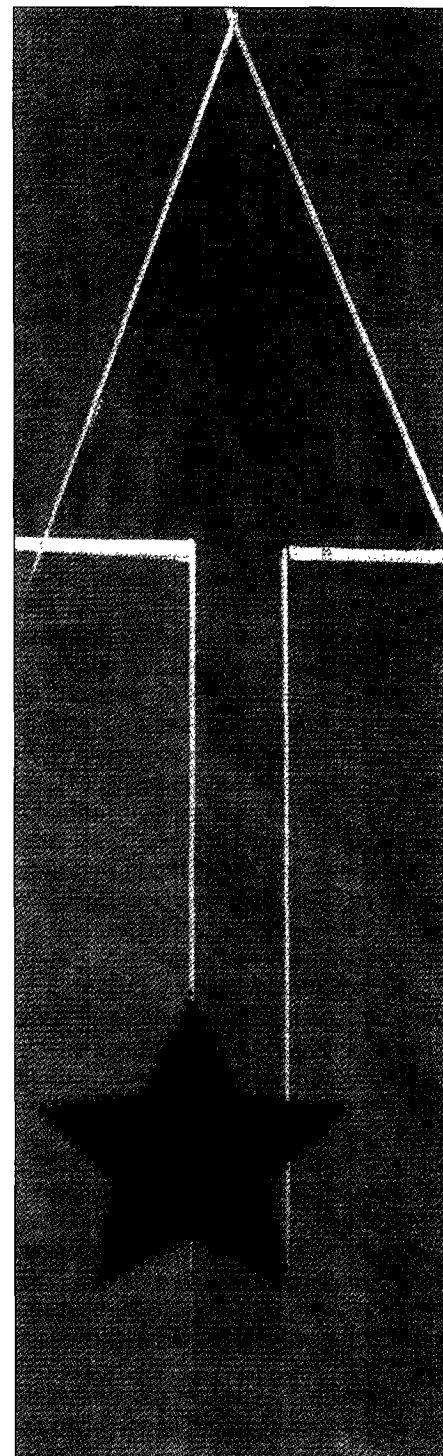
BY ERIC A. HURLEY, ANNE CHAMBERLAIN, ROBERT E. SLAVIN,
AND NANCY A. MADDEN

SUCCESS for All (SFA) is a comprehensive reform model for elementary schools, especially schools that qualify for Title I schoolwide funding. Begun in 1987 in one inner-city school in Baltimore, Success for All is now operating in more than 1,800 schools in 49 states, as well as in five foreign countries. In fall 2000, about one million children were in schools implementing Success for All.

An important strength of Success for All — and a key to its rapid growth — is the amount and quality of the research done to evaluate the program.¹ Studies in 12 school districts have compared SFA to matched

control schools on individually administered measures of reading, especially the Woodcock, Durrell, and Gray informal reading inventories.² These studies have reported that, on average, SFA schools have exceeded controls on these measures by approximately 50% of a standard deviation — or, in grade-equivalents, approximately 2.5 months in first grade, increasing to 1.1 years by fifth grade. These effects have been particularly positive for the lowest-achieving students, and the program has had a profound effect on reducing special education placements and retention.³ In addition to research using student-level data and individually administered measures,

ERIC A. HURLEY and ANNE CHAMBERLAIN are research scientists at the Success for All Foundation (SFAF), Baltimore. ROBERT E. SLAVIN is co-director of the Center for Research on the Education of Students Placed at Risk, Johns Hopkins University, Baltimore, and chairman of SFAF. NANCY A. MADDEN is president and CEO of SFAF. The research reported here was supported by grants from the Office of Educational Research and Improvement, U.S. Department of Education (No. R-117-40005). However, the opinions expressed are those of the authors.



several studies have found positive effects of Success for All on a variety of standardized and norm-referenced measures used in state accountability programs.¹

In all, studies of thousands of elementary school children in dozens of schools and in a variety of settings have been published in the most selective journals in education, and they have consistently reported positive impacts of SFA on student achievement. A few studies have also reported atypical program impacts on student achievement. The balance of evidence would suggest that such results may be attributable to the quality of implementation, a variable known to be critical to all school reform efforts and one that is too often left unconsidered.

Nonetheless, the rigor and variety of evidence indicating positive impacts of SFA on student achievement have established the effectiveness of the SFA program design well beyond the empirical standards typical of educational research. One review of 24 comprehensive school reform designs, conducted by the American Institutes of Research, found very few that had ever been compared to a control group.² Only two, Success for All and Direct Instruction, met the highest standards for the rigor of their evaluations. A later compendium by the Thomas Fordham Foundation came to the same conclusion.³

Despite this research, some scholars have questioned the basic effectiveness of Success for All and the evidence that supports it, charging in particular that the majority of the studies were completed by Johns Hopkins researchers or their colleagues and that "independent studies" have found few benefits beyond the early grades, especially on the group-administered measures for which schools are increasingly held accountable.

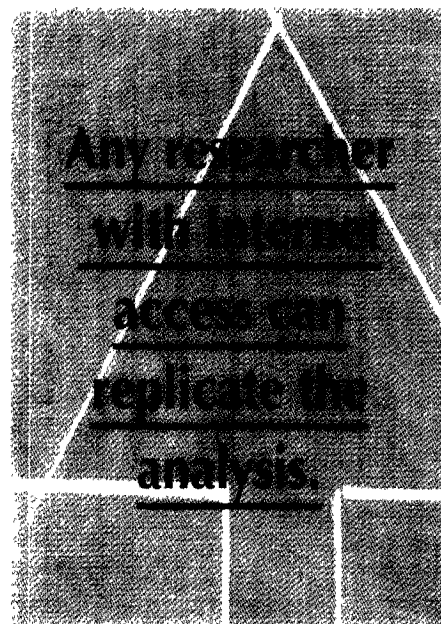
In this article, we use data (available on the Internet) from the Texas Assessment of Academic Skills (TAAS). Such outcome data are often reported by program developers, but usually they have little scientific validity, since the data presented are for schools selected because they happened to have done well in a given year. Yet these standardized test scores are outcomes of greatest importance to educators. The present study deals with this problem of selection bias by including every school that began SFA from 1994 to 1997, a total of 111 schools, and comparing their gains from the year before SFA to 1998. By also including every other school in Texas,

we reduce selection bias further, and we ensure objectivity by the fact that any researcher with Internet access can replicate the analysis. Universal inclusion creates an analysis that is high in generalizability and meaning for policy.

In addition, there is a need to examine the effects of SFA for students of a variety of ethnicities. Most previous investigations of SFA have involved schools that are almost entirely African American or Hispanic, so the results for different ethnic groups have not typically been reported. One study that did report data by ethnic group found that, while SFA increased the achievement of both African American and white students, the gains for African American students were larger relative to control groups. At the end of the study there were no ethnic group differences in the SFA schools, while ethnic group differences remained substantial in control schools. This study involved just two SFA and two control schools, however, and its analyses of ethnicity by treatment interactions have not been repeated.⁴

Positive effects of SFA have been documented many times for African American students, for Hispanic students, and for white students, but there have not been large enough samples of each ethnic group to permit conclusive comparisons of relative impacts for each ethnic group.⁵ Impacts for African American and Hispanic students are particularly important, of course, because these groups typically do not score as well as white students on measures of reading. These differences have profound implications for later school achievement and life success and are a leading factor in the creation of broader social inequities. Convincing evidence that SFA can significantly help to narrow the reading gap between African American, Hispanic, and white students would represent an extremely important development for education policy and practice.⁶ Our study allows for comparisons by ethnic group so that we can illuminate this issue.

We present analyses of data from TAAS reading measures. We sought to evaluate program outcomes in all the 111 Texas schools that began using SFA between 1994 and 1997. TAAS reading scores are collected in all elementary grades, starting in grade 3, so the analyses presented here evaluate the effects of SFA in the upper-elementary grades. Our analysis is by far the largest evaluation of SFA (in fact, it is the largest

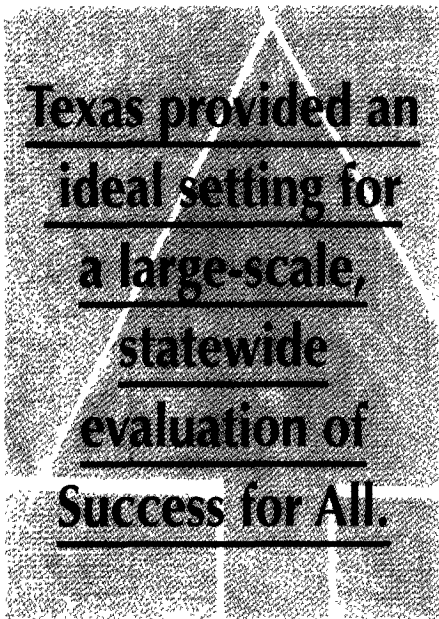


evaluation of any comprehensive reform model ever conducted), and it is the first large-scale study to examine results separately by student ethnicity. Furthermore, while the analysis presented could be considered "first party," the data were taken from the Texas Education Agency (TEA) website and are thus available for replication by any interested parties.

What Is Success for All?

Success for All is a program designed to ensure the reading success of every child by applying a combination of innovative instructional approaches in prekindergarten through grades 5 or 6, one-on-one tutoring for children in the primary grades who are struggling with reading, family support programs, and forms of school organization and professional development that are intended to lead to high-quality implementation of all program elements as well as replicability in a wide variety of circumstances.¹² The program is disseminated by the Success for All Foundation, a nonprofit organization that was spun off from Johns Hopkins University in 1998.

Each element of SFA was originally included because it had a strong base in research, and the program is continually revised to reflect new developments in research as well as lessons learned from the experiences of schools that are implementing the program. A hallmark of SFA is specificity; well-structured student materials, teacher manuals, assessments, and other sup-



ports have been developed for every aspect of the program at every grade level. During the spring before the program is begun at a school, school staffs are encouraged to learn about SFA, to read the research, and to visit other schools using the program. Ultimately, at least 80% of the entire staff must vote to adopt the program. In addition, the school must be able to afford the program, which costs (for a school of 500 students) about \$75,000 to \$80,000 in the first year, \$30,000 in the second year, and \$20,000 in the third year. Moreover, the school must be able to afford sufficient program staff members, who are usually reallocated from other functions.

A child's progress through SFA begins in prekindergarten or kindergarten. The program at this level uses thematic, interdisciplinary units, storytelling and retelling, language development activities, phonemic awareness and alphabet activities, and other elements designed to build children's oral language, background knowledge, concepts of print, and preliteracy skills. From about midway through kindergarten into first grade, the program provides a balanced approach to reading instruction that uses a systematic approach to phonics, taught in the context of meaningful text, very much in line with the recent blue-ribbon reviews of research on the components of effective early literacy instruction.¹³ Children experiencing difficulty with early reading instruction receive daily one-on-one tutoring from certified teachers or from well-trained paraprofessionals in an effort to ensure that each

of them will get off to a good start in reading.

Beyond the first-grade level, teachers use strategies adapted from Cooperative Integrated Reading and Composition.¹⁴ In this approach, children work in small, cooperative learning teams on activities focusing on main idea, summarization, vocabulary building, home reading, and creative writing.

Success for All places much emphasis on family support programs. And the program offers the school's staff strategies for increasing parent involvement, increasing attendance, improving classroom management, preventing behavior problems, integrating social and health services, and solving other nonacademic problems.

A full-time SFA facilitator, usually an experienced teacher from the school's own staff, works with all teachers to ensure effective program implementation and reviews schoolwide assessments administered every eight weeks to help focus attention on children who are not making adequate progress. Trainers from the Success for All Foundation provide schools with about 24 person-days of onsite training and follow-up activity during the first year of implementation, with continuing training and follow-up activity in later years. In addition, there is a weeklong training session for principals and facilitators before the program begins, and conferences for the staffs of experienced schools each spring.

For schools with many students who are learning English, SFA provides two alternatives. Schools with Spanish bilingual or dual-language approaches use a Spanish adaptation, *Exito para Todos*, while those using an ESL (English as a second language) approach use strategies developed to help English language learners build reading skills and English skills in tandem.

The Texas Evaluation

Texas provided an ideal setting for a large-scale, statewide evaluation of Success for All. First, there are a large number of SFA schools in the state; until 1999-2000, Texas had the largest number of SFA schools of any state (it was edged out by California in that academic year). Second, Texas was one of the first states to make its state assessment and school-level demographic data available on the Internet, thus making analyses far easier for researchers. Third,

Texas used the same assessment, with few changes in the test itself or in the basis used to compute test scores, from 1994 to 1998, thereby making longitudinal analyses possible. Unfortunately for researchers, the Texas Education Agency changed test procedures in 1999 and began requiring the inclusion of special education students and significantly changing the rules regarding the administration of the English TAAS to children who were just learning English, especially those children who had been taught reading primarily in Spanish. These changes made the data for 1999 and 2000 not comparable with the earlier data.

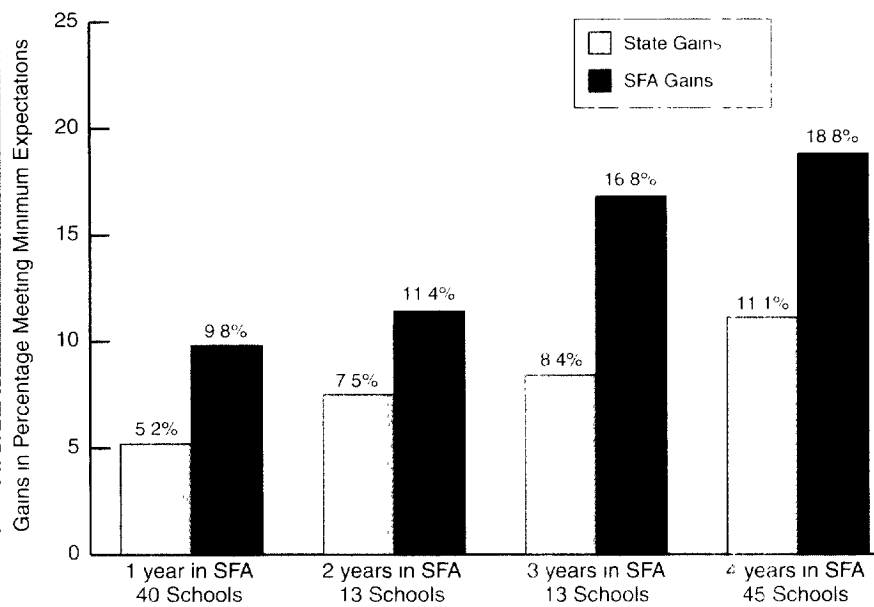
One possible problem with the TAAS is that scores on this measure, expressed as the percentage of students who meet a proficiency standard, rose significantly statewide between 1994 and 1997, and some concern has been expressed that there is a ceiling effect for high achievers. We have tried to address this issue in our analyses.

The 111 Texas schools involved in this evaluation are almost all Title I schoolwide projects and have very high poverty levels. On average, 85% of the children in these schools are designated economically disadvantaged according to state criteria. (The figure for the state as a whole is 45%.) Otherwise, the SFA schools are quite diverse. While most are in large cities (such as Houston, San Antonio, Galveston, and El Paso), many are in small towns or rural areas (such as Eagle Pass, Harlingen, Morton, and Muleshoe), and some are in inner-suburban districts (such as Aldine near Houston and the North East, Northside, and Southside districts near San Antonio). The schools we studied serve much higher percentages of minority students than is true of most Texas schools. According to state data, students in the SFA schools were 25% African American, 62% Hispanic, and 13% white. Corresponding figures for the state are 14% African American, 35% Hispanic, and 47% white. Students with limited proficiency in English made up 27% of the SFA sample, but only 12% of the state's students.

Our Findings

Overall analyses. For an initial, overall analysis, TAAS reading scores, obtained from the TEA website, were averaged for each of the 111 SFA schools across grades 3, 4, and 5. Each cohort consisted of all schools that began to implement SFA in

FIGURE 1.
Gains on TAAS Reading from Preimplementation
Year to 1998, SFA Schools Versus State of Texas,
All Students, Grades 3-5



also noted earlier, SFA schools are far more impoverished than other Texas schools and contain a much higher proportion of minority students and English learners. Because of these differences, the gain scores shown in Figure 1 compare dissimilar populations and should be interpreted with caution.

To create more comparable samples and to obtain information valuable in its own right, we carried out analyses within ethnic groups. Data by ethnic group (African American, Hispanic, and white) were also available on the TEA website. Scores for a given school are reported by TEA if there are at least five students of a given ethnic group in a given grade.

Results for African American students. Results for African American students are shown in Figures 2 and 3. Figure 2 shows gain scores for each of four cohorts (1994, 1995, 1996, and 1997). Our analyses indicated that African American students in 66 SFA schools gained 5.62 percentage points more than those in control schools, and the difference was statistically significant. The effect size was +0.37. Figure 3 shows the same data in raw percentage passing. In this analysis, the African American

the designated academic year (e.g., schools in the 1994 cohort have used the program for four years). Gains from the spring before program inception to the spring of 1998 were computed for each cohort of schools and were compared to gains for the state as a whole over the same period. The gain scores are displayed in Figure 1.

The gains for each cohort of SFA schools were greater than gains for the state during the same time period, with both the raw amount of gain and the relative advantage for SFA schools generally increasing with each additional year of implementation, from a difference of 4.6 percentage points for one-year schools to a difference of 7.7 percentage points for four-year schools. Gain scores were combined across cohorts by subtracting from each SFA school's gain the gain made in the state over the same period. On average, the SFA schools gained 5.85 percentage points more than the state, a difference that was statistically significant. The effect size, the mean difference divided by the standard deviation of school means, was +0.59.

As noted earlier, the relative gains for Success for All schools may have been influenced by a ceiling effect on TAAS. As

FIGURE 2.
Gains on TAAS Reading from Preimplementation
Year to 1998, SFA Schools Versus State of Texas,
African American Students, Grades 3-5

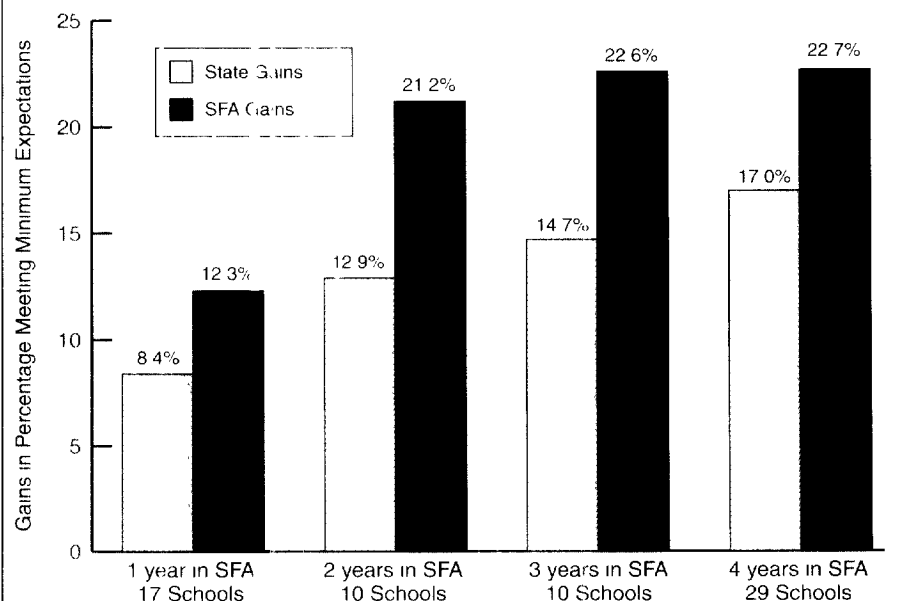
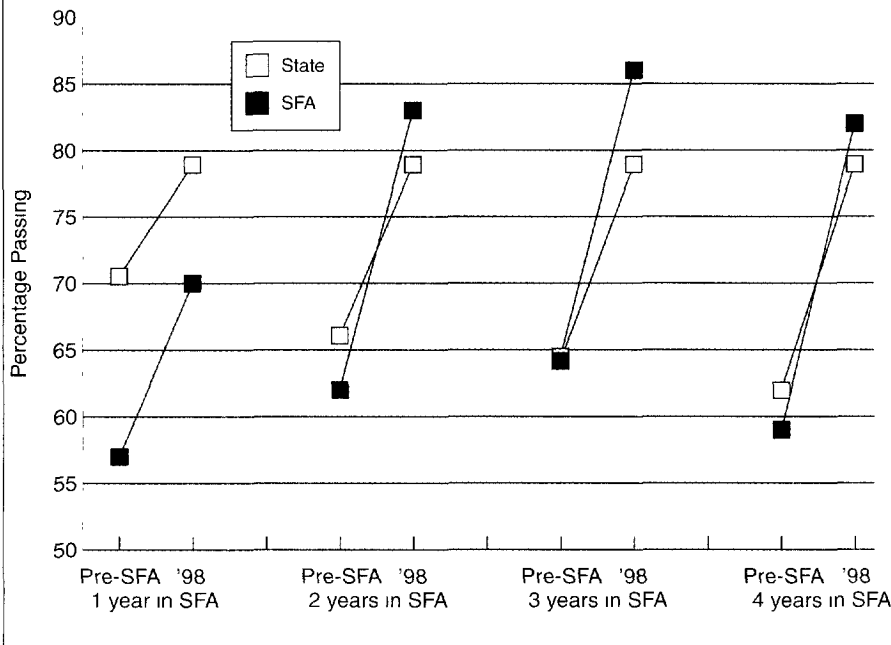


FIGURE 3.
Percentage Passing TAAS Reading from Preimplementation
Year to 1998, SFA Schools Versus State of Texas,
African American Students, Grades 3-5



term of results for Hispanic students, the largest ethnic group in Texas SFA schools, was generally similar to that of African American students. Figures 4 and 5 summarize these results. Not surprisingly, the greatest gains for Hispanic students in SFA schools (relative to those in the state as a whole) were for the 1994 cohort (four years in SFA). In the 1995 cohort, Hispanic students gained somewhat more in the state as a whole. Still, combining across all 95 schools with adequate numbers of Hispanic students, gains for Hispanics in SFA schools were significantly larger than those for Hispanics in other schools. The effect size was +0.28.

Figure 5 shows the same data in raw percentage passing. In this analysis, the Hispanic students are compared to Hispanic students in the state as a whole. And here again there is little evidence of ceiling effects. Unfortunately, there is no way to do separate analyses for Spanish-dominant students taught in Spanish or English, as these data are mixed in with data for English-dominant students of Hispanic origin. There have been other studies of Success for All that did make these distinctions.¹⁵

Results for white students. For white stu-

students are comparable to African American students in the state as a whole, and there is little evidence of ceiling effects. For example, the 1995 cohort (three years in SFA) started at 63% passing and increased to 86% passing, while African American students in the state as a whole started at 64% and increased to 79%. Only in the 1997 cohort (one year in SFA) was there a substantial difference on the pretest between African American students in SFA and those in all of Texas.

The net effect of the gains for African American students in Success for All schools was to substantially narrow the gap between their scores and those of white students. At the pretest, African American students in the 1995 cohort trailed white students by 24.6 percentage points, while at the post-test (1998), they were just 6.5 percentage points behind white students. Other African Americans throughout the state remained 13.8 points behind. The closing of the black/white achievement gap for all Texas students may well be influenced by a ceiling effect for whites, but the *relative* advantage for African Americans in SFA schools is unlikely to be affected by this factor.

Results for Hispanic students. The pat-

FIGURE 4.
Gains on TAAS Reading from Preimplementation
Year to 1998, SFA Schools Versus State of Texas,
Hispanic Students, Grades 3-5

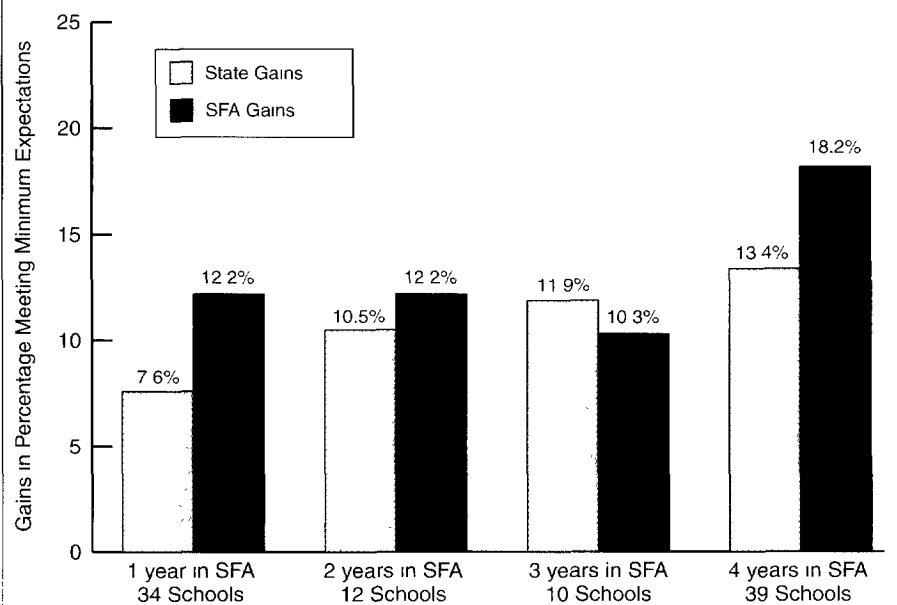
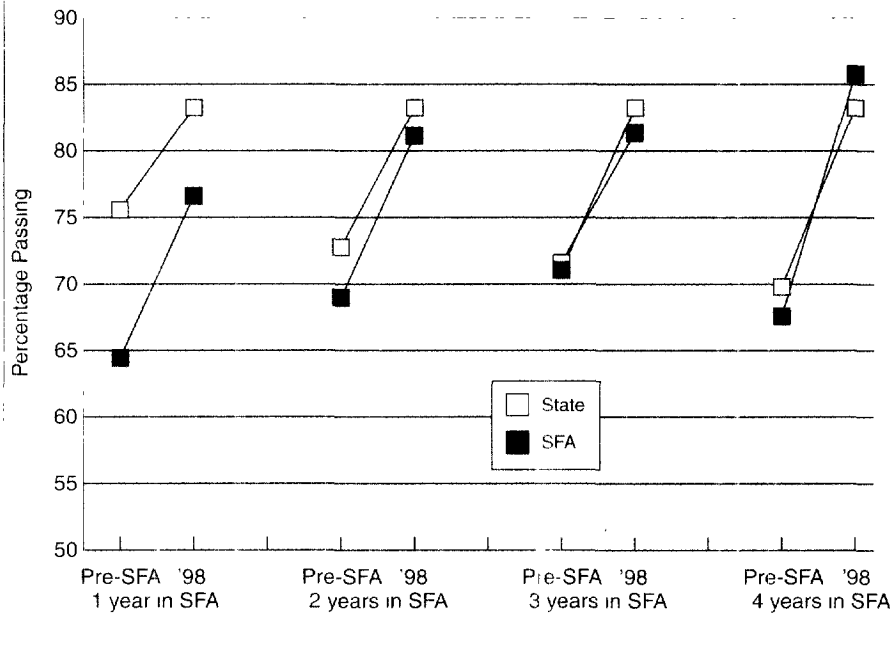


FIGURE 5.
Percentage Passing TAAS Reading from Preimplementation
Year to 1998, SFA Schools Versus State of Texas,
Hispanic Students, Grades 3-5



and Hispanic students, for whom the gap with white students closed significantly more than it did for African American and Hispanic students in other schools throughout the state.

The Texas data are important in several ways. First, they provide evidence from widely available test scores of greater growth on a state accountability measure by SFA schools than was achieved by other schools in the state. The sample of SFA schools, 111 schools serving more than 60,000 children, was large enough to allow for analyses using individual school means as the unit of analysis, which provides for a very conservative test. Clearly, some schools did better than others in both implementation and outcome, but no schools were excluded from the analysis because of poor implementation or for any other reason.

At the policy level, it is crucial to know that whatever differences there may have been among schools, the overall effort produced meaningful gains on measures that all Texas schools are trying to influence. These data suffer from all the random factors that plague all measurement for accountability, such as the effects of mobility, differential special education assign-

ments, the ceiling effect on TAAS scores was most acute. In 1998, 92.7% of white students statewide passed TAAS, an increase from the 85% who passed in 1994. This means that there were many schools that started and ended the period under study with all or nearly all of their white students passing. Fifty of the 111 SFA schools had enough white students to report data for this subgroup. White students in SFA schools gained substantially more than other white students in the 1995 cohort (three years in SFA) and the 1996 cohort (two years in SFA), but the gains were about equal in the larger 1997 and 1994 cohorts. The overall comparison of white students' scores in these 50 schools did not reveal significant differences, but the change was directionally positive, with an effect size of +0.17.

Conclusion

The statewide data for Texas reported here show that Success for All schools are significantly and substantially closing the gap in TAAS reading performance between themselves and the far less impoverished schools in the rest of the state. We found this to be particularly true for African American

FIGURE 6.
Gains on TAAS Reading from Preimplementation
Year to 1998, SFA Schools Versus State of Texas,
White Students, Grades 3-5

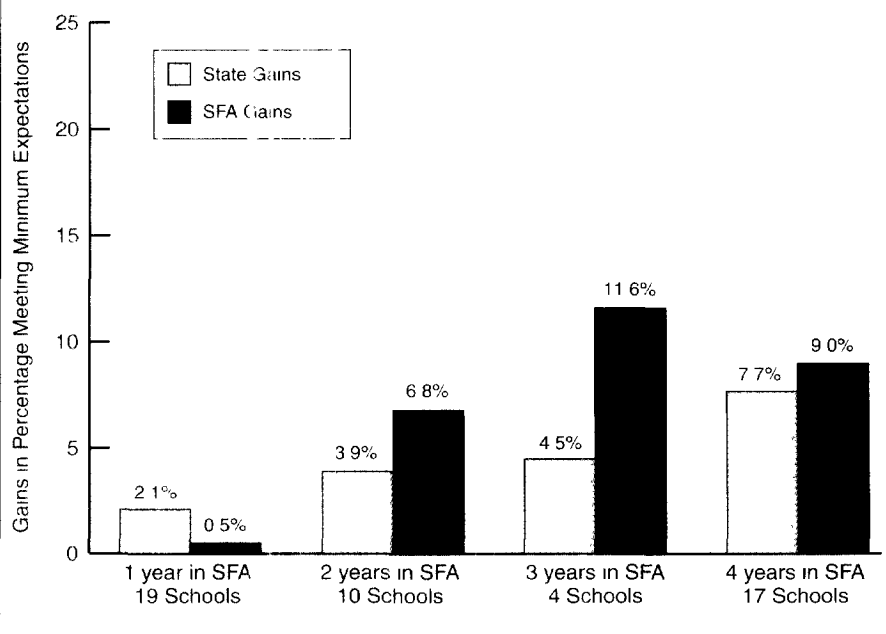
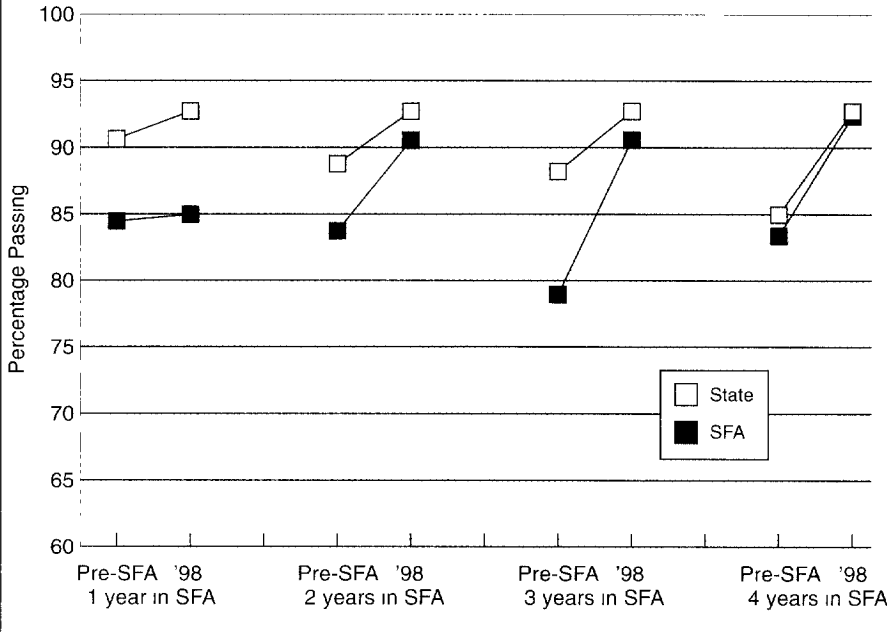


FIGURE 7.
Percentage Passing TAAS Reading from Preimplementation Year to 1998, SFA Schools Versus State of Texas, White Students, Grades 3-5



ments, different amounts of missing data, different test-preparation strategies, and different practices regarding testing of English learners. Yet in a large-scale assessment, these confounding factors tend to balance out and show a true treatment effect.

Further, the findings of this study show again that the impact of Success for All can be seen on standardized test scores in the upper-elementary grades (3-5), not just in the early-elementary grades, where the impacts have historically been documented and are undisputed by critics.

The results presented here also raise some questions for further research. The outcomes for African American and Hispanic students largely mirror the positive results seen in earlier studies. However, there have been many fewer studies that include significant numbers of white students. Those that have been done have found positive effects of SFA for white students or for samples that are majority white.¹⁶ But a study in Ft. Wayne, Indiana, found (as did our study) that the effects were greatest for African Americans and could significantly narrow the achievement gap with whites.¹⁷

For Hispanics as well, we found that

the outcomes of Success for All were very positive, whether the students have been English-dominant, initially Spanish-dominant and taught in Spanish, or initially Spanish-dominant and taught in English.¹⁸

Clearly, more research is needed to examine different effects of Success for All for different ethnic groups, not only to determine what these differences are, but also to understand why they occur. There is no more important policy question in American education than the question of how to eliminate the gap in school performance between African American and Hispanic students and their white counterparts. This difference appears very early in reading performance.¹⁹ Research on Success for All holds out the possibility that programs designed to ensure the reading success of all children may be particularly beneficial for African American and Hispanic children, who have historically been underserved in American schools. And this possibility must be explored further.

1. Rebecca Herman, *An Educator's Guide to Schoolwide Reform* (Washington, D.C.: American Institute for Research, 1999).

2. Robert E. Slavin and Nancy A. Madden, eds., *One*

Million Children. Success for All (Mahwah, N.J.: Erlbaum, 2001)

3. Robert E. Slavin, "Neverstreaming Preventing Learning Disabilities," *Educational Leadership*, February 1996, pp. 4-7; and Lana J. Smith, Steven M. Ross, and Jason P. Casey, *Special Education Analyses for Success for All in Four Cities* (Memphis: Center for Research in Educational Policy, University of Memphis, 1994)

4. Nancy A. Madden et al., "Success for All: Longitudinal Effects of a Restructuring Program for Inner-City Elementary Schools," *American Educational Research Journal*, vol. 30, 1993, pp. 123-48. John Nunnery et al., "An Assessment of Success for All Program Component Configuration Effects on the Reading Achievement of At-Risk First-Grade Students," paper presented at the annual meeting of the American Educational Research Association, New York, 1996, and Slavin and Madden, op cit

5. Herman, op cit

6. James Traub, *Better by Design? A Consumer's Guide to Schoolwide Reform* (Washington, D.C.: Thomas Fordham Foundation, 1999)

7. Steven Ross, Lana Smith, and Jason P. Casey, *1994-95 Success for All Program in Ft. Wayne, IN: Final Report* (Memphis: Center for Research in Educational Policy, University of Memphis, 1995)

8. Ibid.

9. Robert E. Slavin and Nancy A. Madden, "Research on Achievement Outcomes of Success for All: A Summary and Response to Critics," *Phi Delta Kappan*, September 2000, pp. 38-40, 59-66, and idem, *One Million Children*.

10. Patricia Donahue et al., *NAEP 1998 Reading Report Card for the Nation* (Washington, D.C.: National Center for Education Statistics, 1999)

11. Ross, Smith, and Casey, op cit

12. Slavin and Madden, *One Million Children*.

13. National Reading Panel, *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction* (Rockville, Md.: National Institute of Child Health and Human Development, 2000), and Catherine E. Snow, M. Susan Burns, and Peg Griffin, eds., *Preventing Reading Difficulties in Young Children* (Washington, D.C.: National Academy Press, 1998).

14. Robert J. Stevens et al., "Cooperative Integrated Reading and Composition Two Field Experiments," *Reading Research Quarterly*, vol. 22, 1987, pp. 433-54

15. Robert E. Slavin and Nancy A. Madden, "Effects of Bilingual and English as a Second Language Adaptations of Success for All on the Reading Achievement of Students Acquiring English," *Journal of Education for Students Placed at Risk*, vol. 4, 1999, pp. 393-416

16. Steven M. Ross, John Nunnery, and Lana J. Smith, *Evaluation of Title I Reading Programs Amphitheater Public Schools Year 1 1995-96* (Memphis: Center for Research in Educational Policy, University of Memphis, 1996), and Steven M. Ross et al., "Increasing the Academic Success of Disadvantaged Children: An Examination of Alternative Early Intervention Programs," *American Educational Research Journal*, vol. 32, 1995, pp. 773-800

17. Ross, Smith, and Casey, op cit.

18. Slavin and Madden, "Effects of Bilingual and English as a Second Language Adaptations"

19. Donahue et al., op cit