SUCCESS FOR ALL
Summary of Research on Achievement Outcomes

Robert E. Slavin
Johns Hopkins University

Nancy A. Madden
Success for All Foundation

Updated September, 2012

Portions of this report are adapted from Slavin, R.E., Madden, N.A., Chambers, B., & Haxby, B. (2009). *Two million children: Success for All.* Thousand Oaks, CA: Corwin.

An on-line version of this report is available on the Success for All Foundation web site at www.successforall.org/Research/Research-Archive.

ABSTRACT

The purpose of this review is to describe the current state of research on the achievement outcomes of Success for All. Success for All is based on the findings of research on effective instruction for students at risk to direct all aspects of school and classroom organization toward the goal of preventing academic deficits from appearing in the first place, searching out and intensively intervening with any deficits that do appear, and providing students with a rich and full curriculum to enable them to build critical thinking on their firm foundation in basic skills. The commitment of Success for All is to do whatever it takes to see that all children become skilled, strategic, and enthusiastic learners as they progress through the elementary and middle grades.

The results of evaluations of hundreds of Success for All schools in districts in all parts of the United States clearly show that the program increases student reading performance. More than 50 experimental-control comparison studies done by researchers from many institutions, including a national randomized evaluation, have shown positive effects on a wide variety of reading measures. Significant effects have not been seen on every measure at every grade level, but the consistent direction and magnitude of the effects show unequivocal benefits for Success for All students.

AUTHORS' NOTE

Success for All is currently developed and disseminated by the Success for All Foundation, a not-for-profit organization created in 1998. For further information on Success for All, including information on program adoption, contact:

Success for All Foundation

200 West Towsontown Boulevard

Baltimore, MD 21204

1-800-548-4998

FAX: 410-324-4440

WEB SITE: www.successforall.org

E-MAIL: sfainfo@successforall.org

TABLE OF CONTENTS

Despite the constant public outcry about the crisis in American education, every community has one or more outstanding and often widely recognized public schools. Some of these appear to succeed because they serve children of wealthy, well-educated parents, or because they are magnet schools that can screen out unmotivated or low-achieving students. However, there are also schools that serve disadvantaged and minority children in inner city or rural locations and, year after year, produce outstanding achievement outcomes. Such schools play a crucial role in reminding us that the problems of our school system have little to do with the capabilities of children; they provide our best evidence that all children can learn. Yet the success of these lighthouse schools does not spread very far. Excellence can be demonstrated in many individual schools but rarely in whole districts or communities. An outstanding elementary school benefits about 500 children, on average. Yet there are millions of children who are placed at risk by ineffective responses to such factors as economic disadvantage, limited English proficiency, or learning difficulties. How can we make excellence the norm rather than the exception, especially in schools serving many at-risk children? How can effective practices based on research and on the experiences of outstanding schools be effectively implemented every day by hundreds of thousands of teachers?

Success for All is one answer to these questions. Born in one Baltimore school in 1987, Success for All has been adopted by more than 1,000 schools in 47 states, the United Kingdom, and Canada. More than two million children have attended Success for All schools. These schools are highly diverse. They are in most of the largest urban districts, but also hundreds of rural districts, inner suburban districts, and Indian reservations. Most

are Title I schoolwide projects with many children qualifying for free lunches, but many are in much less impoverished circumstances.

Success for All is by far the largest research-based, whole-school reform strategy ever to exist. It is the first model to demonstrate that techniques shown to be effective in rigorous research can be replicated on a substantial scale with fidelity and continued effectiveness. Both the research and the dissemination of Success for All pose an inescapable challenge to educational policy. If replicable excellence is possible, then how can we accept the abysmal performance of so many children? This is not to say that every school needs to adopt Success for All, but what it does imply is that every school needs to create or adopt some program that is no less effective than Success for All. It is unconscionable to continue using ineffective practices if effective ones are readily available and are capable of serving any school that is prepared to dedicate itself to quality implementation.

Success for All: The Promise and the Plan

To understand the concepts behind Success for All, let's start with Ms. Martin's kindergarten class. Ms. Martin has some of the brightest, happiest, and most optimistic kids you'll ever meet. Students in her class are glad to be in school, proud of their accomplishments, certain that they will succeed at whatever the school has to offer. Every one of them is a natural scientist, a storyteller, a creative thinker, a curious seeker of knowledge. Ms. Martin's class could be anywhere, in suburb or ghetto, small town or barrio, it doesn't matter. Kindergartners everywhere are just as bright, enthusiastic, and

confident as her kids are.

Only a few years from now, many of these same children will have lost the spark they all started with. Some will have failed a grade. Some will be in special education. Some will be in long term remediation, such as Title I or other remedial programs. Some will be bored or anxious or unmotivated. Many will see school as a chore rather than a pleasure and will no longer expect to excel. In a very brief span of time, Ms. Martin's children will have defined themselves as successes or failures in school. All too often, only a few will still have a sense of excitement and positive self-expectations about learning. We cannot predict very well which of Ms. Martin's students will succeed and which will fail, but we can predict based on the past that if nothing changes, far too many will fail. This is especially true if Ms. Martin's kindergarten happens to be located in a high-poverty neighborhood, in which there are typically fewer resources in the school to provide top-quality instruction to every child, fewer forms of rescue if children run into academic difficulties, and fewer supports for learning at home. Preventable failures occur in all schools, but in high poverty schools failure can be endemic--so widespread that it makes it difficult to treat each child at risk of failure as a person of value in need of emergency assistance to get back on track. Instead, many such schools do their best to provide the greatest benefit to the greatest number of children possible, but have an unfortunately well-founded expectation that a certain percentage of students will fall by the wayside during the elementary years.

Any discussion of school reform should begin with Ms. Martin's kindergartners. The first goal of reform should be to ensure that every child, regardless of home

background, home language, or learning style, achieves the success that he or she so confidently expected in kindergarten, that all children maintain their motivation, enthusiasm, and optimism because they are objectively succeeding at the school's tasks. Any reform that does less than this is hollow and self-defeating.

What does it mean to succeed in the elementary grades? The elementary school's definition of success, and therefore the parents' and children's definition as well, is overwhelmingly success in reading. Very few children who are reading adequately are retained, assigned to special education, or given long-term remedial services. Other subjects are important, of course, but reading and language arts form the core of what school success means in the early grades.

When a child fails to read well in the early grades, he or she begins a downward progression. In first grade, some children begin to notice that they are not reading adequately. They may fail first grade or be assigned to long term remediation. As they proceed through the elementary and middle grades, many students begin to see that they are failing at their full-time jobs. When this happens, things begin to unravel. Failing students begin to have poor motivation and poor self-expectations, which lead to continued poor achievement, in a declining spiral that ultimately leads to despair, delinquency, and dropout.

Remediating learning deficits after they are already well established is extremely difficult. Children who have already failed to learn to read, for example, are now anxious about reading, and doubt their ability to learn it. Their motivation to read may be low. They may ultimately learn to read but it will always be a chore, not a pleasure. Clearly, the

time to provide additional help to children who are at risk is early, when children are still motivated and confident and when any learning deficits are relatively small and remediable. The most important goal in educational programming for students at risk of school failure is to try to make certain that we do not squander the greatest resource we have: the enthusiasm and positive self-expectations of young children themselves.

In practical terms, what this perspective implies is that schools, and especially Title I, special education, and other services for at-risk children, must be shifted from an emphasis on remediation to an emphasis on prevention and early intervention. Prevention means providing developmentally appropriate preschool and kindergarten programs so that students will enter first grade ready to succeed, and it means providing regular classroom teachers with effective instructional programs, curricula, and professional development to enable them to ensure that most students are successful the first time they are taught. Early intervention means that supplementary instructional services are provided early in students' schooling and that they are intensive enough to bring at-risk students quickly to a level at which they can profit from good quality classroom instruction.

Success for All is built around the idea that every child can and must succeed in the early grades, no matter what this takes. The idea behind the program is to use everything we know about effective instruction for students at risk to direct all aspects of school and classroom organization toward the goal of preventing academic deficits from appearing in the first place; recognizing and intensively intervening with any deficits that do appear; and providing students with a rich and full curriculum to enable them to build on their firm foundation in basic skills. The commitment of Success for All is to do whatever it takes to

see that every child becomes a skilled, strategic, and enthusiastic reader by the end of the elementary grades and beyond.

Usual practices in elementary schools do not support the principle of prevention and early intervention. Most provide a pretty good kindergarten, a pretty good first grade, and so on. Starting in first grade, a certain number of students begin to fall behind, and over the course of time these students are assigned to remedial programs (such as Title I) or to special education, or are simply retained.

Our society's tacit assumption is that those students who fall by the wayside are defective in some way. Perhaps they have learning disabilities, or low IQ's, or poor motivation, or parents who are unsupportive of school learning, or other problems. We assume that since most students do succeed with standard pretty good instruction in the early grades, there must be something wrong with those who don't.

Success for All is built around a completely different set of assumptions. The most important assumption is that every child can learn. We mean this not as wishful thinking or just a slogan, but as a practical, attainable reality. In particular, every child without organic retardation can learn to read. Some children need more help than others and may need different approaches than those needed by others, but one way or another every child can become a successful reader.

The first requirement for the success of every child is prevention. This means providing excellent preschool and kindergarten programs, improving curriculum, instruction, and classroom management throughout the grades, assessing students frequently to make sure they are making adequate progress, and establishing cooperative

relationships with parents so they can support students learning at home.

Top-quality curriculum and instruction from age four on will ensure the success of most students, but not all of them. The next requirement for the success of all students is intensive early intervention. This means one-to-one tutoring for primary-grade students having reading problems. It means being able to work with parents and social service agencies to be sure that all students attend school, have medical services or eyeglasses if they need them, have help with behavior problems, and so on.

The most important idea in Success for All is that the school must relentlessly stick with every child until that child is succeeding. If prevention is not enough the child may need tutoring. If this is not enough he or she may need help with behavior or attendance or eyeglasses. If this is not enough he may need a modified approach to reading or other subjects. The school does not merely provide services to children, it constantly assesses the results of the services it provides and keeps varying or adding services until every child is successful.

Origins of Success for All

The development of the Success for All program began in 1986 as a response to a challenge made to our group at Johns Hopkins University by Baltimore's superintendent, Alice Pinderhughes, its school-board president, Robert Embry, and a former Maryland Secretary of Human Resources, Kalman "Buzzy" Hettleman. They asked us what it would take to ensure the success of every child in schools serving large numbers of disadvantaged students.

At the time, we were working on a book called Effective Programs for Students at Risk (Slavin, Karweit, & Madden, 1989), so we were very interested in this question. After many discussions, the superintendent asked us to go to the next step. We met for months with a planning committee, and finally produced a plan and selected a school to serve as a site. We began in September, 1987 in a school in which all students were African-American and approximately 83% qualified for free lunch.

The first-year results were very positive (see Slavin, Madden, Karweit, Livermon, & Dolan, 1990). In comparison to matched control students, Success for All students had much higher reading scores, and retentions and special education placements were substantially reduced.

In 1988-89, Success for All was expanded in Baltimore to a total of five schools. We also began implementation of Success for All at one of the poorest schools in Philadelphia, in which a majority of the students were Cambodian. This school gave us our first experience in adapting Success for All to meet the needs of limited English proficient students. In 1990-91 we developed a Spanish version of the Success for All beginning reading program, called *Lee Conmigo*, and began to work in more bilingual schools as well as schools providing English as a Second Language instruction (Slavin & Madden, 1999; Cheung & Slavin, in press). In 1992, we received a grant from the New American Schools Development Corporation (NASDC) to add math, science, and social studies to the reading and writing programs of Success for All (Slavin & Madden, 2000), and to help build an organization capable of scaling the program up to serve many more schools.

During the 1990's, Success for All grew exponentially, adding from 40% to 100%

each year from 1989 to 2001. Since 2001, growth has continued at a steady pace. As noted earlier, to date Success for All has been adopted by more than 1,000 schools in 47 states, the United Kingdom, and Canada. The districts range from some of the largest in the country, to small rural districts, including schools on Indian reservations. Success for All reading curricula in Spanish have been developed and researched and are used in bilingual programs throughout the U.S. Most Success for All schools are high-poverty Title I schools, and the great majority are schoolwide projects. Otherwise, the schools vary widely.

Overview of Success for All Components

Success for All is a whole school model that addresses instruction, particularly in reading, as well as schoolwide issues related to leadership, attendance, school climate, behavior management, parent involvement and health that support student achievement (see Slavin, Madden, Chambers, & Haxby, 2009 for more detail).

Literacy Instruction

Learning to read and write effectively is essential for success in school, so Success for All provides in depth support for reading acquisition. Instructional practices, teacher's guides, student materials, assessments and job-embedded professional development are combined to create a powerful reading program.

The Success for All reading program is based on research and effective practices in beginning reading (e.g., Adams, 1990), the National Literacy Panel (2000) and an

appropriate use of cooperative learning to enhance motivation, engagement, and opportunities for cognitive rehearsal. (Slavin, 1995; Stevens, Madden, Slavin, and Farnish, 1987).

Students in grades one and up are regrouped for reading. The students are assigned to heterogeneous, age-grouped classes most of the day, but during a regular 90-minute reading period they are regrouped by reading performance levels into reading classes of students all at the same level. For example, a reading class taught at the 2-1 level might contain first, second, and third grade students all reading at the same level. The reading classes are smaller than homerooms because tutors and other certificated staff (such as librarians or art teachers) teach reading during this common reading period.

Regrouping allows teachers to teach the whole reading class without having to break the class into reading groups. This greatly reduces the time spent in seatwork and increases direct instruction time, eliminating workbooks, dittos, or other follow-up activities which are needed in classes that have multiple reading groups. The regrouping is a form of the Joplin Plan, which has been found to increase reading achievement in the elementary grades (Slavin, 1987).

Reading teachers at every grade level begin the reading time by reading children's literature to students and engaging them in a discussion of the story to enhance their understanding of the story, listening and speaking vocabulary, and knowledge of story structure. In kindergarten and first grade, the program emphasizes development of basic language skills with the use of Story Telling and Retelling (STaR), which involves the students in listening to, retelling, and dramatizing children's literature. Big books as well

as oral and written composing activities allow students to develop concepts of print as they also develop knowledge of story structure. Specific oral language experiences are used to further develop receptive and expressive language.

KinderCorner (SFAF 2012a) offers a full-day theme-based kindergarten program designed to support the development of oral language and vocabulary, early literacy, and social and emotional skills needed for long term success. KinderCorner focuses on getting students talking using cooperative discussion with an integrated set of activities. Opportunities for imaginative play increase both self-regulation and language. Media-based phonemic awareness and early phonics ease students into reading.

Reading Roots (SFAF, 2003) is introduced in first grade. This K-1 beginning reading program uses as its base a series of phonetically regular but meaningful and interesting minibooks and emphasizes repeated oral reading to partners as well as to the teacher. The minibooks begin with a set of "shared stories," in which part of a story is written in small type (read by the teacher) and part is written in large type (read by the students). The student portion uses a phonetically controlled vocabulary. Taken together, the teacher and student portions create interesting, worthwhile stories. Over time, the teacher portion diminishes and the student portion lengthens, until students are reading the entire book. This scaffolding allows students to read interesting literature when they only have a few letter sounds.

Letters and letter sounds are introduced in an active, engaging set of activities that begins with oral language and moves into written symbols. Individual sounds are integrated into a context of words, sentences and stories. Instruction is provided in story

structure, specific comprehension skills, metacognitive strategies for self-assessment and self-correction, and integration of reading and writing. Brief video segments use animations to reinforce letter sounds, puppet skits to model sound blending, and live action skits to introduce key vocabulary.

Spanish bilingual programs use an adaptation of *Reading Roots* called *Lee Conmigo* ("Read With Me"). *Lee Conmigo* uses the same instructional strategies as *Reading Roots*, but is built around shared stories written in Spanish.

When students reach the second grade reading level, they use a program called *Reading Wings* (SFAF, 2011), an adaptation of Cooperative Integrated Reading and Composition (CIRC) (Stevens, Madden, Slavin, & Farnish, 1987). *Reading Wings* uses cooperative learning activities built around story structure, prediction, summarization, vocabulary building, decoding practice, and story-related writing. Students engage in partner reading and structured discussion of stories or novels, and work toward mastery of the vocabulary and content of the story in teams. Story-related writing is also shared within teams. Cooperative learning both increases students' motivation and engages students in cognitive activities known to contribute to reading comprehension, such as elaboration, summarization, and rephrasing (see Slavin, 1995). Research on CIRC has found it to significantly increase students' reading comprehension and language skills (Stevens et al., 1987).

In addition to these story-related activities, teachers provide direct instruction in reading comprehension skills, and students practice these skills in their teams. Classroom libraries of trade books at students' reading levels are provided for each teacher, and

students read books of their choice for homework for 20 minutes each night. Home readings are shared via presentations, summaries, puppet shows, and other formats twice a week during "book club" sessions.

Materials to support *Reading Wings* through the sixth grade level (and beyond) are built around children's literature and around the most widely used basal series and anthologies. Supportive materials have been developed for more than 100 children's novels and for most current basal series (e.g., Houghton Mifflin, Scott Foresman, Harcourt, Macmillan, and Open Court). The upper-elementary Spanish program, *Alas para Leer*, is built around Spanish-language novels and basal series.

Beginning in the second semester of program implementation, Success for All schools usually implement a writing/ language arts program based primarily on cooperative learning principles (SFAF, 2008).

Quarterly Reading Assessments

Four times each year, reading teachers assess student progress through the reading program. The results of the assessments are used to monitor student achievement growth and identify opportunities for acceleration, determine who is to receive tutoring, to suggest other adaptations in students' programs, and to identify students who need other types of assistance, such as family interventions or screening for vision and hearing problems. These data are reviewed schoolwide each quarter to document the effectiveness of the previous quarter's plans, celebrate success, and plan next steps. A computerized data management system called Member Center is used to help school staff organize, manage,

and share this information.

Reading Tutors

A critical element of the Success for All model is the use of tutors to promote students' success in reading. One-to-one tutoring is the most effective form of instruction known (see Slavin, Lake, Davis, & Madden, 2011). Most tutors are certified teachers with experience teaching Title 1, special education, and/or primary reading. Often, well-qualified paraprofessionals also tutor children with less severe reading problems. Tutors work one-on-one with students who are having difficulties keeping up with their reading groups. The tutoring occurs in 20-minute sessions during times other than reading or math periods.

In general, tutors support students' success in the regular reading curriculum, rather than teaching different objectives. For example, the tutor generally works with a student on the same story and concepts being read and taught in the regular reading class. However, tutors seek to identify learning problems and use different strategies to teach the same skills. They also teach metacognitive skills beyond those taught in the classroom program.

Initial decisions about reading group placement and the need for tutoring are based on informal reading inventories that the tutors give to each child. Subsequent reading group placements and tutoring assignments are made based on curriculum-based assessments given every eight weeks, which include teacher judgments as well as more formal assessments. First graders receive priority for tutoring, on the assumption that the primary function of the tutors is to help all students be successful in reading the first time,

before they fail and become remedial readers. Two forms of computer-assisted tutoring are now used in most Success for All schools. One, called Alphie's Alley (Chambers et al., 2008), provides computer assistance to tutors working one-on-one with children. The other, Team Alphie (Chambers et al., 2011a) allows a tutor to serve up to six children, working in pairs. Schools often use Team Alphie as a first attempt to solve reading problems, and then use Alphie's Alley (or non-technology one-to-one instruction) for children who still need more help, in a response-to-intervention (RTI) model (Chambers et al., 2011b).

Preschool and Kindergarten

Most Success for All schools provide a half-day preschool and/or a full-day kindergarten for eligible students. The preschool and kindergarten programs (SFAF, 2001; SFAF 2012) focus on providing a balanced and developmentally appropriate learning experience for young children. The curriculum emphasizes the development and use of language. It provides a balance of academic readiness and non-academic music, art, and movement activities in a series of thematic units. Readiness activities include use of language development activities and Story Telling and Retelling (STaR), in which students retell stories read by the teachers. Reading instruction is phased in during kindergarten.

Leading for Success

Insuring success for every child requires relentlessness. Schools must have systems that enable them to assess needs, set goals for improvement, make detailed plans to implement effective strategies, and monitor progress on a child by child basis. In Success

for All, the tool that guides this schoolwide collaboration is called Leading for Success (SFAF, 2012b).

Leading for Success is built around a distributive leadership model, and engages all school staff in a network of teams that address key areas targeted for continuous improvement. The leadership team manages the Leading for Success process and convenes the staff at the beginning of the school year and at the end of each quarter to assess progress, and set goals and agendas for next steps. Staff members participate in different teams to address areas of focus that involve schoolwide supports for students and families as well as support for improving implementation of instructional strategies to increase success (SFAF, 2012c).

Schoolwide Solutions Teams

Attendance: The attendance team is responsible for insuring that all students are in school on time each day, and for putting solutions in place if attendance or tardiness problems occur at individual, subgroup, grade or school levels.

Parent and Family Involvement: The Parent and Family Involvement Team first works towards good relations with parents and to increase involvement in the school. Team members organize "welcome" visits for new families, opportunities for informal chats among parents and school staff members, workshops for parents on supporting achievement and general parenting issues, and volunteer opportunities.

Intervention: Students struggling with learning or behavioral issues are offered additional supports through a case management approach to intervention when preventive

strategies are not enough.

Cooperative Culture: One group of school staff members addresses the development of a positive school culture. A set of lessons focused on cooperation and self-regulation, Getting Along Together, is taught by all teachers. All school staff support, implement the conflict resolution and problem solving skills presented throughout the building and throughout the day.

Community Connections: This group of staff members builds links to community resources to meet family and student needs. Based on a needs assessment, this group may seek resources for health screenings, eyeglass procurement, volunteer listeners, incentives to be used to recognize success, or other opportunities.

Instructional Component Teams

In Success for All, teachers support one another to develop strong implementations of the research-proven instructional practices embodied in classroom resources and to consider specific student needs. Instructional component teams for KinderCorner, Reading Roots, Reading Wings, Reading Edge and tutoring meet biweekly to share data, discuss strategies, and identify implementation targets.

Each of the teams described above prepares a document every grading period that tracks student progress in the areas for which they are responsible and document goals and strategies for continuous improvement. The Leading for Success process builds shared accountability for success and deep support for continued growth across the school.

Program Facilitator

A program facilitator works at each school to oversee (with the principal) the operation of the Success for All model. The facilitator helps plan the Success for All program, helps the principal with scheduling, and visits classes and tutoring sessions frequently to help teachers and tutors with individual problems. He or she works directly with the teachers on implementation of the curriculum, classroom management, and other issues, helps teachers and tutors deal with any behavior problems or other special problems, and coordinates the activities of the Family Support Team with those of the instructional staff.

Teachers and Teacher Training

The teachers and tutors are regular certified teachers. They receive detailed teacher's manuals supplemented by three days of inservice at the beginning of the school year, followed by classroom observations and coaching throughout the year. For classroom teachers of grades one and above and for reading tutors, training sessions focus on implementation of the reading program (either Reading Roots or Reading Wings), and their detailed teachers' manuals cover general teaching strategies as well as specific lessons. Preschool and kindergarten teachers and aides are trained in strategies appropriate to their students' preschool and kindergarten models. Tutors later receive two additional days of training on tutoring strategies and reading assessment.

Throughout the year, additional inservice presentations are made by the facilitators and other project staff on such topics as classroom management, instructional pace, and

cooperative learning. Facilitators also organize many informal sessions to allow teachers to share problems and problem solutions, suggest changes, and discuss individual children. The staff development model used in Success for All emphasizes relatively brief initial training using a great deal of video and simulations, with extensive classroom follow-up, coaching, and group discussion.

Special Education

Every effort is made to deal with students' learning problems within the context of the regular classroom, as supplemented by tutors. Tutors evaluate students' strengths and weaknesses and develop strategies to teach in the most effective way. In some schools, special education teachers work as tutors and reading teachers with students identified as learning disabled as well as other students experiencing learning problems who are at risk for special education placement. One major goal of Success for All is to keep students with learning problems out of special education if at all possible, and to serve any students who do qualify for special education in a way that does not disrupt their regular classroom experience (see Slavin, 1996; Borman & Hewes, 2002).

Relentlessness

While the particular elements of Success for All may vary from school to school, there is one feature we try to make consistent in all: A relentless focus on the success of every child. It would be entirely possible to have tutoring, curriculum change, family support, and other services, yet still not ensure the success of at-risk children. Success does

not come from piling on additional services but from coordinating human resources around a well-defined goal, constantly assessing progress toward that goal, and never giving up until success is achieved.

None of the elements of Success for All are completely new or unique. All are based on well-established principles of learning and rigorous instructional research. What is most distinctive about them is their schoolwide, coordinated, and proactive plan for translating positive expectations into concrete success for all children. Every child can complete elementary school a confident, strategic, and joyful learner and can maintain the enthusiasm and positive self-expectations they had when they came to first grade. The purpose of Success for All is to see that this vision can become a practical reality in every school.

Research on Success for All

One of the guiding principles in the development of Success for All is an emphasis on rigorous evaluation. The elements of the program are themselves derived from current research on reading and writing, on early childhood, second language learning, and special education, and on parent involvement, professional development, and school change, among many others. However, it is not enough for a program to be based on good research: it must also be rigorously and repeatedly evaluated in many schools over meaningful periods of time in comparison to similar control schools.

Success for All is arguably the most extensively evaluated whole-school reform strategy ever to exist. It was originally conceived, developed, and evaluated at Johns

Hopkins University within a research center, currently called the Center for Research and Reform in Education, with federal funding that required extensive and rigorous evaluation. Its implementation has been thoroughly evaluated and proven to be effective for student achievement in numerous research studies, most of which were done by independent researchers. In 2010, Success for All received the highest score in the Department of Education's Investing in Innovation (i3) scale-up competition, which required strong, scientific evidence of effectiveness. The i3 award included funding for a multi-year third-party randomized evaluation of Success for All by MDRC, the most recent large-scale study of the approach.

Independent Reviews

A number of independent reviews of research on whole-school reform strategies and reading programs have all concluded that Success for All is among the most successfully evaluated of programs. In a 2005 report by the Comprehensive School Reform Quality Center at the American Institutes for Research (CSRQ, 2005), Success for All was cited as one of only two elementary comprehensive designs that met the highest standards for research given in a review of 22 programs. The CSRQ review identified 31 "convincing" studies of Success for All, 10 for Direct Instruction, and no more than 6 for any other program. The same conclusion was reached in an earlier AIR review (Herman, 1999) and in studies commissioned by the Fordham Foundation (Traub, 1999) and the Milliken Family Foundation (Schacter, 1999). A meta-analysis of research on 29 comprehensive school reforms by Borman, Hewes, Overman, & Brown (2003) listed

Success for All among three CSR models with "strongest evidence of effectiveness." Social Programs that Work (www.evidencebasedprograms.org), which uses particularly stringent standards, listed Success for All in grades K-2 as its only proven school or classroom reform model.

Reviews of Research by the Center for Research and Reform in Education

The Center for Research and Reform in education has completed multiple reviews of research on reading programs, including elementary reading, middle and high schools reading, struggling readers, and reading for English language learners. These reviews show that Success for All has been proven through rigorous research to help students succeed. In the elementary reading review, Success for All was one of only three programs found to have "strong evidence of effectiveness" for beginning readers (Slavin, Lake, Chambers, Cheung, & Davis, 2009). Success for All was also found to have "strong evidence of effectiveness" for struggling readers (Slavin, Lake, Davis, & Madden, 2010).

For middle and high school reading, Success for All's "Reading Edge" program was found to have "moderate evidence of effectiveness" (Slavin, Cheung, Groff, and Lake, 2008). The same was true for the most recent English language learners review, in which Success for All was the only program reviewed to receive a rating above "limited evidence of effectiveness" (Cheung & Slavin, 2012).

Longitudinal Studies

Longitudinal experiments evaluating SFA have been carried out since the earliest

program implementations in Baltimore and Philadelphia. Later, third-party evaluators at the University of Memphis, Steven Ross, Lana Smith, and their colleagues, added evaluations in many districts across the U.S. Studies focusing on English language learners in California have been conducted by researchers at WestEd, a federally-funded regional educational laboratory. Each of these evaluations compared Success for All schools to comparison schools on measures of reading performance, starting with cohorts in kindergarten or in first grade and following these students as long as possible (details of the evaluation design appear below). Several studies were able to follow Success for All schools for many years. Data comparing matched SFA and traditional control schools on individual reading measures have been collected from schools in many U.S. districts, and other studies have compared Success for All to a variety of alternative reform models, have compared full and partial implementations of SFA, and have made other comparisons. In 2006, a three-year national randomized experiment involving 41 schools was completed that compared SFA and control schools. In addition, there have been many studies involving group-administered standardized tests including both national norm-referenced tests and state criterion-referenced tests used in state accountability programs. Experimental-control comparisons have also been carried out in Canada, England, Australia, and Israel.

The largest number of studies has compared the achievement of students in Success for All schools to that of children in matched comparison schools using traditional methods, including locally-developed Title I reforms.

A common evaluation design, with variations due to local circumstances, was used

in a foundational set of Success for All evaluations carried out by researchers at Johns Hopkins University, the University of Memphis, and WestEd. Each Success for All school was matched with a control school that was similar in poverty level (percent of students qualifying for free lunch), historical achievement level, ethnicity, and other factors. Schools were also matched on district-administered standardized test scores given in kindergarten or on Peabody Picture Vocabulary Test (PPVT) scores given by the evaluators in the fall of kindergarten or first grade.

The measures used in these evaluations were as follows:

1. Woodcock Reading Mastery Test. Three Woodcock scales, Word Identification, Word Attack, and Passage Comprehension, were individually administered to students by trained testers. Word Identification assesses recognition of common sight words, Word Attack assesses phonetic synthesis skills, and Passage Comprehension assesses comprehension in context. Students in Spanish bilingual programs were given the Spanish versions of these scales.
2. Durrell Analysis of Reading Difficulty. The Durrell Oral Reading scale was also individually administered to students in grades 1-3 in some studies. It presents a series of graded reading passages which students read aloud, followed by comprehension questions.
3. Gray Oral Reading Test. Comprehension and passage scores from the Gray Oral Reading Test were obtained in some studies from students in grades 4-5.

Analyses of covariance with pretests as covariates were used to compare raw scores in all evaluations, and separate analyses were conducted for students in general and, in many studies, for students in the lowest 25% of their grades at pretest.

Each of the evaluations summarized in this chapter follows children who began in Success for All in first grade or earlier, in comparison to children who had attended the control school over the same period. Students who start in it after first grade are not considered to have received the full treatment (although they are of course served within the schools).

Reading Outcomes

National Randomized Evaluation of Success for All

The definitive evaluation of the reading outcomes of Success for All was a U.S. Department of Education-funded evaluation (Borman et al., 2005 a, b; Borman, Slavin, Cheung, Chamberlain, Madden, & Chambers, 2007) involving 41 Title I schools throughout the U.S. Schools were randomly assigned to use Success for All or to continue with their existing reading programs in grades K-2. At the end of the three-year study, children in the Success for All schools were achieving at significantly higher levels than control students on all three measures, using conservative hierarchical linear modeling analyses with school as the unit of analysis. In effect sizes (difference in adjusted posttests divided by unadjusted standard deviations), the differences were ES=+0.38 for Word Attack, ES=+0.23 for Word Identification, and ES=+0.21 for Passage Comprehension. This study is of particular importance for several reasons. First, the use of random assignment to

conditions eliminates selection bias (the possibility that schools that chose SFA might have been better schools than the control schools). Random assignment has become extremely important in program evaluation as the U.S. Department of Education has virtually required randomized designs and emphasized this design element in its What Works Clearinghouse reviews of effective programs. Second, the large sample size allowed for the use of hierarchical linear modeling (HLM), which uses the school as the unit of analysis. This is the appropriate analysis for schoolwide interventions (all previous studies used the student as the unit of analysis). Third, the size of the evaluation is of great importance for the study's policy impact. In small studies, there is always the possibility that researchers can ensure high-quality implementations. Cronbach et al. (1980) called this the "superrealization" of a program's impact, where the program is evaluated at a level of quality far beyond what could be achieved at a large scale. In the Borman et al. (2007) study, implementation of SFA was actually found to be of lower quality than that typical of SFA schools. A study on such a large scale is a good representation of the likely policy effect, or what would be expected if a district or state implemented the program on a broad scale. For example, the Borman et al. (2007) findings imply that if many schools serving African-American or Hispanic students experienced Success for All, the minority-White achievement gap (about 50% of a standard deviation on the National Assessment of Educational Progress) would be reduced by half.

Matched Longitudinal Studies

In the 1990's, researchers at Johns Hopkins University and other research

institutions carried out a series of longitudinal matched studies to evaluate Success for All. A common design and set of measures was used for these studies, as noted above. Schools in 13 districts across the country were involved.

The results of the matched experiments evaluating Success for All are summarized in Figure 1 for each grade level, 1-5, and for followup measures into grades 6 and 7. The analyses average cohort means for experimental and control schools. A cohort is all students at a given grade level in a given year. For example, the Grade 1 graph compares 68 experimental to 68 control cohorts, with cohort (50-150 students) as the unit of analysis. In other words, each bar is a mean of scores from more than 6000 students. Grade equivalents are based on the means, and are only presented for their informational value. No analyses were done using grade equivalents.

Combining across studies, statistically significant (p=.05 or better) positive effects of Success for All (compared to controls) were found on every measure at every grade level, 1-5 (Slavin & Madden, 1993). For students in general, effect sizes averaged around a half standard deviation at all grade levels. Effects were somewhat higher than this for the Woodcock Word Attack scale in first and second grades, but in grades 3-5 effect sizes were more or less equivalent on all aspects of reading. Consistently, effect sizes for students in the lowest 25% of their grades were particularly positive, ranging from ES=+1.03 in first grade to ES=+1.68 in fourth grade. Again, cohort-level analyses found statistically significant differences favoring low achievers in Success for All on every measure at every grade level. A followup study of Baltimore schools found that positive program effects continued into grade 6 (ES=+0.54) and grade 7 (ES=+0.42), when students were in middle

schools.

==========

Figure 1 Here

==========

Long-Term Outcomes

Borman & Hewes (2002) carried out a longitudinal study of children from five Success for All and five control schools in Baltimore. They located children citywide who had attended these schools in first grade and remained from one to six years (mean = 3.8 years). At posttest, students who had been promoted each year would have been in the eighth grade, having been out of a Success for All school for at least three years.

Long-term differences were found on achievement, retentions, and special education placements. In achievement, former SFA students still scored significantly better than controls on standardized, district-administered CTBS reading scores (ES = +0.29, $p<.001$). Surprisingly, there was also a small difference favoring the former SFA students in math (ES=+0.11, $p<.05$), even though mathematics was not part of the intervention.

Success for All students were far less likely to have been retained in elementary school. Nine percent of SFA students and 23% of control students had been retained at least once by fifth grade (ES=+0.27, $p<.001$). Similarly, control students spent 50% more time in special education, on average, than SFA students (ES=+0.18, $p<.001$).

The importance of the Borman & Hewes (2002) study is in its finding that at entry to high school, former Success for All students were in much better shape than control

students. The best predictors of high school success are reading achievement and avoidance of retention and special education placements. Success for All students were substantially higher on all of these measures.

Quality and Completeness of Implementation

Not surprisingly, effects of Success for All are strongly related to the quality and completeness of implementation. In a large study in Houston, Nunnery, Slavin, Ross, Smith, Hunter, & Stubbs (1996) found that schools implementing all program components obtained better results (compared to controls) than did schools implementing the program to a moderate or minimal degree.

In this study, 46 school staffs were allowed to select the level of implementation they wanted to achieve. Some adopted the full model, as ordinarily required elsewhere; some adopted a partial model; and some adopted only the reading program, with few if any tutors, and half-time facilitators or no facilitators. Many of the schools used the Spanish bilingual form of SFA and were assessed in Spanish.

==========

Figures 2 & 3 Here

==========

Figures 2 and 3 summarize the results. The figures show effect sizes comparing SFA to control schools on individually administered measures. On the English measure (Figure 1.2), effect sizes were very positive for the schools using the full program (ES=+0.47), less positive for those with a medium degree of implementation (ES=+0.31),

but for those implementing the fewest program elements, effect sizes were slightly negative (ES= -0.13), indicating that the control groups achieved somewhat better scores. Among schools teaching in Spanish, there were too few certified teacher-tutors for any school to qualify as a high implementer (due to a shortage of teachers). However, medium implementers scored very well (ES=+.31), while low implementers scored less well (ES=+.19) (see Figure 3).

A Memphis study (Ross, Smith, Lewis, & Nunnery, 1996; Ross, Smith, & Nunnery, 1998) compared the achievement of eight Success for All schools to that of four schools using other restructuring designs, matched on socioeconomic status and PPVT scores. Each pair of SFA schools had one school rated by observers as a high implementer, and one rated as a low implementer. In the 1996 cohort, first grade results depended entirely on implementation quality. Averaging across the four Woodcock and Durrell scales, every comparison showed high-implementation SFA schools scored higher than their comparison schools, while low-implementation SFA schools scored lower (Ross et al., 1996). However, by second grade, Success for All schools exceeded comparison schools, on average, and there was less of a clear relationship with the original implementation ratings, perhaps because implementation quality changed over the two year period. Similarly, the 1997 first grade cohort did not show a clear pattern with respect to quality of implementation.

Cooper, Slavin, & Madden (1998), in an interview study, found that high-quality implementations of Success for All depended on many factors, including district and principal support, participation in national and local networks, adequacy of resources, and

genuine buy-in at the outset on the part of all teachers.

Effects on District-Administered Standardized Tests

The formal evaluations of Success for All have relied on individually administered assessments of reading. The Woodcock and Durrell scales used in these assessments are far more accurate than district-administered tests, and are much more sensitive to real reading gains. They allow testers to hear children actually reading the material of increasing difficulty and to respond to questions about what they have read. The Woodcock and Durrell are themselves nationally standardized tests, and produce norms (e.g., percentiles, NCE's and grade equivalents) just like any other standardized measure.

However, educators often want to know the effects of innovative programs on the kinds of group administered standardized tests they are usually held accountable for. To obtain this information, researchers have often analyzed standardized or state criterion-referenced test data comparing students in experimental and control schools. The following sections briefly summarize findings from these types of evaluations.

Memphis

One of the most important early independent evaluations of Success for All is a study carried out by researchers at the University of Tennessee-Knoxville for the Tennessee State Department of Education (Ross, Sanders, & Wright, 1998). William Sanders, the architect of the Tennessee Value-Added Assessment System (TVAAS), carried out the analysis. The TVAAS gives each school an expected gain, based primarily on school

poverty levels, and compares it to actual scores on the Tennessee Comprehensive Assessment Program (TCAP). TVAAS scores above 100 indicate gains in excess of expectations; those below 100 indicate the opposite. Sanders compared TVAAS scores in eight Memphis Success for All schools to those in a) matched comparison schools, and b) all Memphis schools.

==========

Figure 4 Here

==========

Figure 4 summarizes the results for all subjects assessed. At pretest, the Success for All schools were lower than both comparison groups on TVAAS. However, after two years, they performed significantly better than comparison schools in reading, language, science, and social studies.

A third-year evaluation found that Success for All schools averaged the greatest gains and highest levels on the TVAAS of six restructuring designs (Co-nect, Accelerated Schools, Audrey Cohen College, ATLAS, and Expeditionary Learning), as well as exceeding controls, averaging across all subjects (Ross, Wang, Sanders, Wright, & Stringfield, 1999).

An earlier study of Success for All schools in Memphis also showed positive effects on the TCAP. This was a longitudinal study of three Success for All and three control schools carried out by Ross, Smith & Casey (1995). On average, Success for All schools

exceeded controls on TCAP reading by an effect size of +0.38 in first grade and +0.45 in second grade.

State of Texas

From 1994 to 2003, the State of Texas administered the Texas Assessment of Academic Success, or TAAS, in grades 3, 4, and 5. It also assessed fourth grade reading in 1993. Texas was one of the first states to put its scores for every school every year on the Internet, making it possible to compare TAAS gains made by Success for All schools throughout the state to gains in the state as a whole. The importance of this is that it makes possible an independent assessment of program outcomes.

Reading data are summarized in Figure 5 (adapted from Hurley, Slavin, Chamberlain, & Madden, 2001). The data, from 117 high-poverty Title I schools using Success for All throughout the state, are organized according to the year the program began, focusing on gains from the spring before program implementation to spring, 1998, averaged across grades 3-5. The gains for Success for All cohorts are compared to average gains in all other schools throughout the state over the same time periods. Because TAAS scores, expressed as the percent of students meeting minimum expectations, had been gradually rising statewide, gains for Success for All and other schools were larger the longer the time period involved. However, Success for All schools consistently gained more than other schools, with the relative advantage increasing with each year in the program. All differences between SFA and state gains were statistically significant ($p<.05$), using school means as the unit of analysis.

==========

Figure 5 Here

==========

In raw scores, rather than gains, what happened is that Success for All schools, far more impoverished than the state average, started far below state averages and then increased to near equality with state averages. For example, the 1993 cohort (grade 4 only, as this was the only grade tested in 1993) began 32.5 percentile points behind the state in percent meeting minimum standards on TAAS reading, but five years later was nearly indistinguishable from the state (88.1% passing for the state vs. 85.0% for 46 SFA schools). Similar patterns were seen for the 1994, 1995, and 1996 cohorts, all of which started far below state means and, as of 1998, were within 1.4 to 3.8 percentile points of state means. The only cohort that did not close the gap to this extent was the latest, the 1997 cohort (35 schools). This cohort started 17.3 percentile points behind the state and, one year later, was 12.8 points behind, a real improvement but not to virtual equality.

Houston, Texas

In Houston, Success for All was begun on a large scale in 1995. Forty-six schools adopted Success for All as part of a study (Nunnery et al., 1996) in which they were allowed to implement either the full program, the reading program only, or something in between. As noted earlier, the full-implementation schools obtained excellent outcomes on individually-administered tests given to subsamples, in comparison to control schools, while moderate-implementation schools obtained less positive outcomes and low-

implementation schools did not differ from controls (recall Figures 2 and 3).

Baltimore, Maryland

A longitudinal study in Baltimore from 1987-1993 collected CTBS scores on the original five Success for All and control schools (Madden, Slavin, Karweit, Dolan, & Wasik, 1993). On average, Success for All schools exceeded control schools at every grade level. The differences were statistically and educationally significant. By fifth grade, Success for All students were performing 75% of a grade equivalent ahead of controls (ES=+0.45) on CTBS Total Reading scores (see Slavin, Madden, Dolan, Wasik, Ross, & Smith, 1994).

Ft. Wayne, Indiana

An evaluation in two schools in Ft. Wayne, Indiana (Ross, Smith, & Casey, 1997) found positive effects of Success for All on the reading comprehension scale of the ISTEP, Indiana's norm-referenced achievement test. In first grade, the effect size was +0.49 for students in general and +1.13 for the lowest-performing 25%. In second grade, effect sizes were +0.64, and in third grade, ES=+.13.

Louisville, Kentucky

A study carried out by the Jefferson County Public Schools (Muñoz & Dossett, 2004) evaluated Success for All in three experimental and three matched control schools in Louisville, KY, over a four-year period. The SFA schools made twice the gains as control

schools in normal curve equivalent scores (ES=+0.11, p<.05). SFA schools also made greater gains in attendance, and substantially reduced out-of-school suspensions. They also had higher teacher ratings of perceptions of educational quality and job satisfaction and higher student ratings of school climate and educational quality.

New York City's Chancellor's District

From 1996 to 2003, the New York City Public Schools established a special city-wide region, called the Chancellor's District, for the very lowest-achieving elementary and middle schools in the city. These schools implemented Success for All in reading, Everyday Mathematics in math, and had significant resources to provide high-quality professional development to teachers. Independent researchers at New York University did an evaluation of the outcomes of the program (Phenix, Siegel, Zaltsman, & Fruchter, 2004). The results indicated significantly greater gains on state English Language Arts scores for the Chancellor's District schools than for matched control schools. There were no differences in mathematics. The Chancellor's District included elements in addition to Success for All, so the study is not purely an SFA evaluation, but the findings do support the use of Success for All as a key component of an intensive approach to accelerating achievement in very difficult urban schools.

Los Angeles

A study by Mason (2005) of five comprehensive school reform models in Los Angeles found that only Success for All significantly improved students' achievement on

the Stanford Achievement Test. Across 8 SFA schools, students who experienced SFA starting in the first or second grade gained significantly more than those in matched control schools in reading and language arts. Students who participated in SFA for all three years of the study made particularly substantial gains compared to controls. As was found in the Borman & Hewes (2002) long-term followup study, there were also small positive effects on mathematics measures, even though math was not part of the SFA intervention. Students who started SFA after second grade did not experience significant effects, however.

International Evaluations of Success for All Adaptations

Several studies have assessed the effects of adaptations of Success for All in countries outside of the U.S. These adaptations have ranged from relatively minor adjustments to accommodate political and funding requirements in Canada and England to more significant adaptations in Mexico, Australia, and Israel.

The Canadian study (Chambers, Abrami, & Morrison, 2001) involved one school in Montreal, which was compared to a matched control school on individually-administered reading measures. Results indicated significantly better reading performance in the Success for All school than in the control school both for special needs students (a large proportion of the SFA students) and for other students. Similarly, a study of five SFA schools in Nottingham, England, found that Success for All students gained more in reading than did students in a previous cohort, before the program was introduced, and gained more on Key Stage 1 (age 7) and Key Stage 2 (age 11) reading tests than did

English schools overall (Harris, Hopkins, Youngman, & Wordsworth, 2001; Hopkins, Youngman, Harris, & Wordsworth, 1999). Later studies by Russ & Harris (2005) also found positive effects of the program on Key Stage 2 performance. Combining data from all 16 SFA schools in England, Slavin, Wordsworth, & Jones-Hill (2005) found that Year 6 students in SFA schools gained 13.4 percentage points on Key Stage 2 from 2001 to 2004, while other schools in England gained 3.0 percentage points.

Because of language and cultural differences, the most significant adaptation of Success for All was made to use the program in Israel with both Hebrew-speaking children in Jewish schools and Arabic-speaking children in Israeli Arab schools, all in or near the northern city of Acre. The implementation involved community interventions focusing on parent involvement, integrated services, and other aspects in addition to the adapted Success for All model. In comparison to control groups, Success for All first graders performed at significantly higher levels on tests of reading and writing (Hertz-Lazarowitz, 2001).

Australian researchers created a substantially simplified adaptation of Success for All, which they called SWELL. SWELL used instructional procedures much like those used in Success for All, but used books adapted for the Australian context. Only the early grades were involved. Schools did not have full-time facilitators or family support programs, and they may or may not have provided any tutoring. Two studies of SWELL found positive effects of the program on reading performance in comparison to control groups and to Reading Recovery schools (Center, Freeman, & Robertson, 2001; Center, Freeman, Mok, & Robertson, 1997).

Finally, a Mexican study in Juarez, near El Paso, Texas, found substantial gain on Spanish reading tests in three schools implementing an adaptation of Success for All (Calderón, 2001).

The international studies of programs adapted from Success for All have importance in themselves, of course, but also demonstrate that the principles on which Success for All are based transfer to other languages, cultures, and political systems. In addition, they provide third-party evaluations of Success for All in diverse contexts, strengthening the research base for Success for All principles and practices.

Success for All in the Middle Grades: The Reading Edge

Two major evaluations of The Reading Edge (SFAF, 2013d) have shown the potential of the program to accelerate student achievement on the reading skills that lead to success in high school. In the first (Slavin, Daniels, and Madden, 2005), seven middle schools around the U.S. that used The Reading Edge were matched with local schools that used traditional textbooks. The differences were striking. The Reading Edge averaged a gain of 24.6% on state reading assessments over a three-year period. Control schools gained only 2.2%, and other schools in the same states averaged a gain of 4.2%.

In the second study, two middle schools in Florida and West Virginia both randomly assigned sixth graders to use The Reading Edge or to continue with their traditional instruction (Slavin, Chamberlain, Daniels, & Madden, 2009). Findings supported a conclusion that the effects of The Reading Edge are modest but reliable in high-poverty middle schools.

Research on the Success for All middle school program has been reviewed by the federally funded What Works Clearinghouse. No middle school program was given a higher rating for research quality and effectiveness.

The Reading Edge High School

An evaluation of The Reading Edge High School was completed in 2011 (Madden et al., 2011). This quasi-experimental study compared students in five high-poverty Philadelphia-area high schools to similar schools in the same area. Controlling for eighth grade scores on state tests, students in The Reading Edge gained significantly more than controls on the overall Group Reading Assessment and Diagnostic Evaluation (GRADE) and on the Vocabulary subtest, but not on Sentence Comprehension or Passage Comprehension. The results suggest that reading performance can be improved in high-poverty urban high schools through an integrated program of instructional materials and professional development in cooperative learning strategies.

Changes in Effect Sizes Over Years of Implementation

One interesting trend in outcomes from comparisons of Success for All and control schools relates to changes in effects sizes according to the number of years a school has been implementing the program. Figure 6, which summarizes these data, was created by pooling effect sizes for all cohorts in their first year of implementation, all in their second year, and so on, regardless of calendar year.

==========

Figure 6 Here

==========

Figure 6 shows that mean reading effect sizes progressively increased with each year of implementation. For example, Success for All first graders scored substantially better than control first graders at the end of the first year of implementation (ES=+0.49). The experimental-control difference was even higher for first graders attending schools in the second year of program implementation (ES=+0.53), increasing to an effect size of +0.73 for schools in their fourth implementation year. A similar pattern was apparent for second and third grade cohorts.

The data summarized in Figure 6 show that while Success for All has an immediate impact on student reading achievement, this impact grows over successive years of implementation. Over time, schools may become increasingly able to provide effective instruction to all of their students, to approach the goal of success for *all.*

Comparisons With Other Programs

A few studies have compared outcomes of Success for All to those of other whole-school reform strategies.

As noted earlier, a study of six restructuring designs in Memphis on the Tennessee Value Added Assessment System (TVAAS) found that Success for All schools had the highest absolute scores and gain scores on the TVAAS, averaging across all subjects (Ross et al., 1999). The TVAAS is a measure that relates performance on the Tennessee

Comprehensive Achievement Test to "expected" performance. The designs, in addition to Success for All, were Co-nect, Accelerated Schools, Audrey Cohen College, ATLAS, and Expeditionary Learning.

Similarly, the Mason (2005) study in Los Angeles, discussed earlier, compared schools implementing five CSR models to matched controls. In addition to SFA, the programs were America's Choice, Co-nect, Different Ways of Knowing, and Urban Learning Centers. Among these, only SFA students who began the program in grades 1 or 2 gained significantly more than controls.

A study in Clover Park, Washington, compared Success for All to Accelerated Schools (Hopfenberg & Levin, 1993), an approach that, like Success for All, emphasizes prevention and acceleration over remediation, but unlike Success for All does not provide specific materials or instructional strategies to achieve its goals. In the first year of the evaluation, the Success for All and Accelerated Schools programs had similar scores on individually administered reading tests and on a writing test (Ross, Alberg, & McNelis, 1997). By second grade, however, Success for All schools were scoring slightly ahead of Accelerated Schools in reading, and significantly ahead in writing (Ross, Alberg, McNelis, & Smith, 1998).

In addition, the Consortium for Policy Research in Education completed a study of comprehensive school reform programs in more than 120 elementary schools that compared the Accelerated Schools Project, America's Choice, and Success for All to control schools (Rowan et al., 2009). The study identified America's Choice and Success for All as being characterized by significant amounts of instructional guidance, which

promoted instructional change. Students in Success for All gained significantly more in reading in grades 1-3 than did those in the other programs (ES=+0.40).

Success for All and English Language Learners

While language of instruction is an essential concern for children who are acquiring English, the *quality* of instruction (and corresponding achievement outcomes) is at least as important, whatever the initial language of instruction may be (Cheung & Slavin, in press). There is a need for better programs for teaching in the home language and then transitioning to English, and for better programs for teaching English language learners in English with support from English as a second language strategies. Both development and research on Success for All have focused on both of these issues.

Six studies have evaluated adaptations of Success for All with language minority children (see Cheung & Slavin, in press; Slavin & Madden, 1999). Three of these evaluated *Éxito Para Todos* ("Success for All" in Spanish), the Spanish bilingual adaptation, and three evaluated a program adaptation incorporating English as a second language strategies.

Bilingual Studies. One study compared students in *Éxito Para Todos* to those in a matched comparison school in which most reading instruction was in English. Both schools served extremely impoverished, primarily Puerto Rican student bodies in inner-city Philadelphia. Not surprisingly, *Éxito Para Todos* students scored far better than control students on Spanish measures. More important was the fact that after transitioning to all-English instruction by third grade, the *Éxito Para Todos* students scored significantly better

than controls on measures of English reading. These differences were significant on Word Attack, but not on Word Identification or Passage Comprehension.

An evaluation of *Éxito Para Todos* in California bilingual schools was reported by Livingston & Flaherty (1997), who studied three successive cohorts of students. On Spanish reading measures, *Éxito Para Todos* students scored substantially higher than controls in first grade (ES= +1.03), second grade (ES= +0.44), and third grade (ES= +.23). However, the second and third grade differences almost certainly understate the true effects; the highest-achieving students in the bilingual programs were transitioned early to English-only instruction, and the transition rate was twice as high in the *Éxito Para Todos* classes as in the controls.

A large study in Houston compared LEP first graders in 20 schools implementing *Éxito Para Todos* to those in 10 control schools (Nunnery, Slavin, Madden, Ross, Smith, Hunter, & Stubbs, 1996). As an experiment, schools were allowed to choose Success for All/*Éxito Para Todos* as it was originally designed, or to implement key components. The analysis compared three levels of implementation: high, medium, and low. None of the *Éxito Para Todos* programs were categorized as "high" in implementation, as a bilingual teacher shortage made it impossible to hire certified teachers as Spanish tutors, a requirement for the "high implementation" designation. Medium-implementation schools significantly exceeded their controls on all measures (mean ES= +0.24). Low implementers exceeded controls on the Spanish Woodcock Word Identification and Word Attack scales, but not on Passage Comprehension (mean ES= +0.17).

One additional study evaluated Bilingual Cooperative Integrated Reading and

Composition (BCIRC), which is closely related to *Alas Para Leer*, the bilingual adaptation of Reading Wings. This study, in El Paso, Texas, found significantly greater reading achievement (compared to controls) for English language learners in grades 3-5 transitioning from Spanish to English reading (Calderón, Hertz-Lazarowitz, & Slavin, 1998).

ESL Studies. Three studies have evaluated the effects of Success for All with English language learners being taught in English. In this adaptation, ESL strategies (such as total physical response) are integrated into instruction for all children, whether or not they are limited in English proficiency. The activities of ESL teachers are closely coordinated with those of other classroom teachers, so that ESL instruction directly supports the Success for All curriculum, and ESL teachers often serve as tutors for LEP children.

The first study of Success for All with English language learners took place in Philadelphia. Students in an Asian (mostly Cambodian) Success for All school were compared to those in a matched school that also served many Cambodian-speaking children. Both schools were extremely impoverished, with nearly all children qualifying for free lunches.

At the end of a six-year longitudinal study, Success for All Asian fourth and fifth graders were performing far ahead of matched controls. On average, they were 2.9 years ahead of controls in fourth grade (median ES= +1.49), and 2.8 years ahead in fifth grade (median ES= +1.33). Success for All Asian students were reading about a full year above grade level in both fourth and fifth grades, while controls were almost two years below

grade level. Non-Asian students also significantly exceeded their controls at all grade levels (see Cheung & Slavin, 2005; Slavin & Madden, 1999).

The California study described earlier (Livingston & Flaherty, 1997) also included many English language learners who were taught in English. Combining results across three cohorts, Spanish-dominant English language learners performed far better on English reading measures in Success for All than in matched control schools in first grade (ES= +1.36) and second grade (ES= +0.46), but not in third grade (ES= +0.09). As in the bilingual evaluation, the problem with the third grade scores is that many high-achieving children were transitioned out of the ESL designation in the Success for All schools, reducing apparent experimental-control differences. Corresponding effect sizes for students who spoke languages other than English or Spanish were +0.40 for first graders, +0.37 for second graders, and +0.05 for third graders.

An Arizona study (Ross, Nunnery, & Smith, 1996) compared Mexican-American English language learners in two urban Success for All schools to those in three schools using locally-developed Title I reform models and one using Reading Recovery. Two SES school strata were compared, one set with 81% of students in poverty and 50% Hispanic students and one with 53% of students in poverty and 27% Hispanic students. Success for All first graders scored higher than controls in both strata. Hispanic students in the high-poverty stratum averaged three months ahead of the controls (1.75 vs. 1.45). Hispanic students in the less impoverished stratum scored slightly above grade level (1.93), about one month ahead of controls (1.83).

The effects of Success for All for language minority students are not statistically

significant on every measure in every study, but the overall impact of the program is clearly positive, both for the Spanish bilingual adaptation, *Éxito Para Todos,* and for the ESL adaptation. What these findings suggest is that whatever the language of instruction may be, student achievement in that language can be substantially enhanced using improved materials, professional development, and other supports.

Success for All and Academically Talented Students

Because it is primarily used in high-poverty schools, Success for All is sometimes seen as a "remedial" program, or one only for students who are at risk. Yet this is not the case. Success for All has been used successfully in all kinds of communities with all kinds of children. The effects of Success for All for high achievers are very positive (Slavin, 2006). A three-year longitudinal study of Success for All by Slavin, Madden, Karweit, Dolan, & Wasik (1992) found that in comparison to control schools, three times as many students in five Baltimore elementary schools were reading two or more years above grade level.

Success for All and Special Education

Perhaps the most important goal of Success for All is to place a floor under the reading achievement of all children, to ensure that every child performs adequately in this critical skill. This goal has major implications for special education. If the program makes a substantial difference in the reading achievement of the lowest achievers, then it should reduce special education referrals and placements. Further, students who have IEP's

indicating learning disabilities or related problems are typically treated the same as other students in Success for All. That is, they receive tutoring if they need it, participate in reading classes appropriate to their reading levels, and spend the rest of the day in age-appropriate, heterogeneous homerooms. Their tutor and/or reading teacher may be a special education teacher, but otherwise students with IEPs are not treated differently. One-to-one tutoring in reading, plus high-quality reading instruction in the mainstream at the student's appropriate level, should be more effective than the small-group instruction provided in special education classes. For this reason we expect that students who have been identified as being in need of special education services will perform substantially better than similar students in traditional special education programs.

The philosophy behind the treatment of special education issues in Success for All is called "neverstreaming" (Slavin, 1996). That is, rather than waiting until students fall far behind, are assigned to special education, and then may be mainstreamed into regular classes, Success for All schools intervene early and intensively with students who are at risk to try to keep them out of the special education system. Once students are far behind, special education services are unlikely to catch them up to age-appropriate levels of performance. Students who have already failed in reading are likely to have an overlay of anxiety, poor motivation, poor behavior, low self-esteem, and ineffective learning strategies that are likely to interfere with learning no matter how good special education services may be. Ensuring that all students succeed in the first place is a far better strategy, if it can be accomplished. In Success for All, the provision of research-based preschool, kindergarten, and first grade reading, one-to-one tutoring, and family support services are likely to give

the most at-risk students a good chance of developing enough reading skills to remain out of special education, or to perform better in special education than would have otherwise been the case.

The data relating to special education outcomes clearly support these expectations. Several studies have focused on questions related to special education. One of the most important outcomes in this area is the consistent finding of particularly large effects of Success for All for students in the lowest 25% of their classes. While effect sizes for students in general have averaged around + 0.50 on individually administered reading measures, effect sizes for the lowest achievers have averaged in the range of +1.00 to +1.50 across the grades. In the longitudinal Baltimore study only 2.2% of third graders averaged two years behind grade level, a usual criterion for special education placement. In contrast, 8.8% of control third graders scored this poorly. Baltimore data also showed a reduction in special education placements for learning disabilities of about half (Slavin et al., 1992). A longitudinal study following Baltimore children to eighth grade found that students who had been in control schools had spent 50% more time in special education, on average, than those who had been in SFA schools (Borman & Hewes, 2002). A study of two Success for All schools in Ft. Wayne, Indiana found that over a two year period 3.2% of Success for All students in grades K-1 and 1-2 were referred to special education for learning disabilities or mild mental handicaps. In contrast, 14.3% of control students were referred in these categories (Smith, Ross, & Casey, 1994).

Taken together, these findings support the conclusion that Success for All both reduces the need for special education services (by raising the reading achievement of very

low achievers) and reduces special education referrals and placements. Both of these outcomes have significant consequences for long-term costs of educating students placed at risk.

Another important question concerns the effects of the program on students who have already been assigned to special education. Smith et al. (1994) combined first grade reading data from special education students in Success for All and control schools in four districts: Memphis, Ft. Wayne (IN), Montgomery (AL), and Caldwell (ID). Success for All special education students scored substantially better than controls (mean ES =+.59).

Embedding Technology in Success for All

The Success for All Foundation has worked to add technology tools to its early reading programs, and studies of these additions find them to be effective in improving students' reading performance. *Reading Reels*, used in kindergarten and first grade classes, provides appealing video content to supplement Reading Roots. This includes animations to teach letter sounds, puppet skits to teach sound blending, and live action skits to teach vocabulary. In a study in which schools in Hartford, Connecticut were randomly assigned to use SFA either with or without *Reading Reels*, students who experienced the videos performed significantly better on the Woodcock Word Attack scale than those who did not experience *Reading Reels* (Chambers, Cheung, Madden, Slavin, & Gifford,2006).

A second technology enhancement is called *Alphie's Alley*. It is designed for use in SFA tutoring. *Alphie's Alley* helps tutors assess their students, plan their instruction, and provide them with compelling, animated presentations and practice opportunities.

Embedded in the content are professional development videos in which experienced tutors demonstrate tutoring strategies. A study involving 25 SFA schools randomly assigned children in tutoring (and their tutors) to tutoring with or without *Alphie's Alley*. In schools that used the program as intended, *Alphie's Alley* students scored significantly better than students given usual paper-and-pencil tutoring on the Woodcock Letter-Word Identification and Word Attack scales as well as DIBELS Fluency (Chambers, Abrami, Tucker, Slavin, Madden, Cheung, & Gifford,2008).

A third randomized study evaluated outcomes in schools that used both *Reading Reels* and *Alphie's Alley*. Students who received tutoring and experienced both embedded technology interventions scored substantially better than tutored SFA students who did not experience the technology on the Woodcock Letter-Word and Word Attack scales and the Gray Oral Reading Test Fluency and Comprehension scales. Students who were not tutored, and therefore experienced *Reading Reels* but not *Alphie's Alley*, also scored better than non-tutored SFA students who did not experience the videos on Woodcock and Gray reading measures (Chambers, Slavin, Madden, Abrami, Tucker, Cheung, & Gifford, 2008).

Another study evaluated the relative effects of computer-assisted tutoring in small groups (*Team Alphie*) and the regular one-to-one tutoring provided to struggling readers in 33 high-poverty Success for All schools. It also compared how efficient *Team Alphie* was compared to traditional one-to-one tutoring in terms of the number of children who received tutoring services. In this year-long study, the lowest–scoring first and second graders in each school were assigned to tutoring. In the *Team Alphie* schools, students were tutored in groups of 6 in 45-minute daily sessions. In the control schools, students were

tutored for 20 minutes daily, using the standard one-to-one tutoring process used in Success for All. Analyses of covariance of students' Woodcock-Johnson III Tests of Achievement reading scores indicated that the first grade treatment group significantly out-performed the one-to-one tutoring group on all three covariate-adjusted reading measures, with no significant differences between the second-grade treatment and control groups. Schools using *Team Alphie* were able to tutor 31% more first graders than the control schools and 46% more second graders. This study shows that a computer-assisted small group tutoring program can be at least as effective as one-to-one tutoring and serve many more struggling readers (Chambers, Slavin, Madden, Abrami, Karanzalis, & Gifford, 2011).

Taken together, these studies suggest that using multimedia content embedded in teachers' lessons along with Success for All can significantly enhance learning for children. This type of application, in which technology supplements instead of replacing teachers' instruction, may help teachers reinforce content and skills through visual as well as auditory pathways. The positive findings have led the Success for All Foundation to include *Reading Reels*, *Alphie's Alley*, and *Team Alphie* as standard components of Reading Roots.

A study of Success for All's Writing Wings with Multimedia (WWM) has also been completed. The program uses a writing process approach with a strong emphasis on cooperative learning, as well as embedded multimedia segments in which humorous skits model components of the writing process, cooperative learning, writing genres, and metacognitive strategies. In a year-long evaluation, 63 teachers were randomly assigned to WWM or control conditions within 22 schools in 11 states, with a total of 922 third and

fourth graders. Students were given writing prompts in October and May. There were no significant differences on HLM analyses, but using ANCOVAs at the student level there were significant differences on a total writing score, and ratings of Style and Mechanics, and no differences on Ideas and Organization. The findings partially support the use of cooperative learning and embedded multimedia to improve outcomes of writing process models (Madden, Slavin, Logan, & Cheung, 2011).

Teachers' Attitudes Toward Success for All

Three studies have examined teachers' attitudes toward Success for All using questionnaires. Ross, Smith, Nunnery, & Sterbin (1995) surveyed teachers involved in six restructuring designs, including Success for All, and found that Success for All schools had the most positive attitudes toward the success of the implementation. However, all designs were rated relatively positively, and there was more variation among schools implementing the same designs than between models.

Rakow & Ross (1997) studied teacher attitudes in five Success for All schools in Little Rock, Arkansas. Once again, responses varied widely from school to school, but overall effects were very positive. For example, 70% of teachers agreed that SFA was having a positive effect in their schools, and 78% felt "positively about using the SFA model."

Datnow & Castellano (2000) also examined teachers' attitudes in extensive observations and interviews in three California schools with substantial Latino majorities. They described 64% of the teachers as "supportive" or "strongly supportive," 28%

"accepting," and 8% (three teachers) "opposed." Among the "accepting" teachers were some teachers who personally did not like the program, but still felt it was working for their children.

Muñoz & Dossett (2004), in a four-year study of six schools in Louisville, Kentucky, found that SFA teachers gave higher ratings than control teachers on measures of educational quality and job satisfaction.

Perhaps the best indicator of teacher support for Success for All is not from a study, but from a vote. In spring, 1999, the San Antonio Independent School District, responding to a severe budget shortfall and a change of superintendents, required teachers in all schools using restructuring designs to vote on whether to keep these designs or to return to the district's program. A vote of 80% in favor was required to keep the program. Across 24 Success for All schools, the average vote in favor was 81% positive. In contrast, votes for the five other designs (37 schools) averaged 36.5% positive.

Conclusion

The results of evaluations of dozens of Success for All schools in districts in all parts of the U.S., Canada, the UK, and other countries clearly show that the program increases student reading performance. A large, national randomized evaluation found clear positive effects of the program, compared to a control group. Across more than fifty matched studies done by dozens of researchers, Success for All students learned significantly more than matched control students. Significant effects were not seen on every measure at every grade level, but the consistent direction and magnitude of the

effects show unequivocal benefits for Success for All students. Effects on district-administered standardized tests and criterion-referenced tests used in state accountability programs reinforce the findings of the studies using individually administered tests. Large impacts have been seen on the achievement of limited English proficient students in both bilingual and ESL programs, and on both reducing special education referrals and improving the achievement of students who have been assigned to special education.

The Success for All evaluations have used reliable and valid measures, in particular individually administered tests that are sensitive to all aspects of reading: comprehension, fluency, word attack, and word identification. Positive effects on state accountability assessments and on other standardized measures have also been documented many times. Performance of Success for All students has been compared to that of students in similar control schools, who provide the best indication of what students without the program would have achieved. Replication of high-quality experiments in such a wide variety of schools and districts is extremely unusual. As noted earlier, reviews of research by the Comprehensive School Reform Quality Center (2005), Borman et al. (2003), the American Institutes for Research (Herman, 1999) and the Fordham Foundation (Traub, 1999) all found Success for All to be one of only two or three whole-school elementary reform strategies to have rigorous, frequently replicated evidence of effectiveness.

An important indicator of the robustness of Success for All is the fact that schools stay with the program. As of 2012, the median school using SFA has done so for 11 years. When schools do drop the program, it is usually due to a district decision (forced by policy changes or funding cuts), not a school decision. Hundreds of Success for All schools have

survived changes of superintendents, principals, facilitators, and other key staff, major cuts in funding, and other serious threats to program maintenance.

The research summarized here demonstrates that comprehensive, systemic school-by-school change can take place on a broad scale in a way that maintains the integrity and effectiveness of the model. The schools we have studied are typical of the larger set of schools currently using Success for All in terms of quality of implementation, resources, demographic characteristics, and other factors. Program outcomes are not limited to the original home of the program. The widely held idea based on the Rand study of innovation (Berman & McLaughlin, 1978; McLaughlin, 1990) that comprehensive school reform must be invented by school staffs themselves is certainly not supported in research on Success for All. While the program is adapted to meet the needs of each school, and while school staffs must agree to implement the program by a vote of 75% or more, Success for All is an externally developed program with specific materials, manuals, and structures. The observation that the program can be implemented and maintained over considerable time periods and can be effective in each of its replication sites certainly supports the idea that every school staff need not reinvent the wheel.

References

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

Berman, P., & McLaughlin, M. (1978). *Federal programs supporting educational change: A model of education change, Vol. VIII: Implementing and sustaining innovations.* Santa Monica, CA: Rand.

Borman, G., & Hewes, G. (2003*).* Long-term effects and cost effectiveness of Success for All. *Educational Evaluation and Policy Analysis, 24* (2), 243-266.

Borman, G., Slavin, R. E., Cheung, A., Chamberlain, A., Madden, N. A., & Chambers, B. (2005). The national randomized field trial of Success for All: Second-year outcomes. *American Educational Research Journal 42 (4), 673-696.*

Borman, G., Slavin, R. E., Cheung, A., Chamberlain, A., Madden, N. A., & Chambers, B. (2007). Final reading outcomes of the national randomized field trial of Success for All. *American Educational Research Journal*, 44(3), 701–731.

Borman, G.D., Hewes, G.M., Overman, L.T., & Brown, S. (2003) Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73* (2), 125-230.

Borman, G.D., Slavin, R.E., Cheung, A., Chamberlain, A., Madden, N., & Chambers, B. (2005). Success for All: First year results from the National Randomized Field Trial. *Educational Evaluation and Policy Analysis, 27* (1), 1-22.

Calderón, M. (2001). Success for All in Mexico. In R.E. Slavin & N.A. Madden (Eds.), *Success for All: Research and reform in elementary education*. Mahwah, NJ:

Erlbaum.

Calderón, M., Hertz-Lazarowitz, R., & Slavin, R.E. (1998). Effects of Bilingual Cooperative Integrated Reading and Composition on students making the transition from Spanish to English reading. *Elementary School Journal, 99,* (2), 153-165.

Center, Y., Freeman, L., & Robertson, G. (2001). A longitudinal evaluation of the Schoolwide Early Language and Literacy Program (SWELL). In R.E. Slavin & N.A. Madden (Eds.), *Success for All: Research and reform in elementary education*. Mahwah, NJ: Erlbaum.

Center, Y., Freeman, L., Mok, M., & Robertson, G. (1997, March). *An evaluation of Schoolwide Early Language and Literacy (SWELL) in six disadvantaged New South Wales schools*. Paper presented at the annual meeting of the American Educational Research Associaton, Chicago.

Chamberlain, A., Daniels, C., Madden, N. A., & Slavin, R. E. (2007). A randomized evaluation of the Success for All middle school reading program. *Middle Grades Reading Journal, 2* (1), 1–22.

Chambers, B., Abrami, P.C., & Morrison, S. (2001). Can Success for All succeed in Canada? In R.E. Slavin (Ed.), *Success for All: Research and reform in elementary education*. Mahwah, NJ: Erlbaum.

Chambers, B., Abrami, P.C., Tucker, B.J., Slavin, R.E., Madden, N.A., Cheung, A., & Gifford, R. (2008). Computer-assisted tutoring in Success for All: Reading outcomes for first grade. *Journal of Research on Effective Education*, 1(2), 120-137.

Chambers, B., Cheung, A., Madden, N., Slavin, R. E., & Gifford, R. (2006). Achievement effects of embedded multimedia in a Success for All reading program. *Journal of Educational Psychology*, 98(1), 232-237.

Chambers, B., Slavin, R. E., Madden, N. A., Abrami, P. C., Karanzalis, M., & Gifford, R. (2011). Small-group computer-assisted tutoring to improve reading outcomes for struggling first and second graders. *Elementary School Journal, 111*(4), 625–640.

Chambers, B., Slavin, R.E., Madden, N.A., Abrami, P.C., Tucker, B.,J., Cheung, A., & Gifford, R. (2008). Technology infusion in Success for All: Reading outcomes for first graders. *Elementary School Journal*, 109(1), 1–15..

Cheung, A. & Slavin, R.E. (in press). Effective reading programs for Spanish dominant English language learners (ELLs) in the elementary grades: A synthesis of research. *Review of Educational Research.*

Cheung, A., & Slavin, R.E. (2005). Effective reading programs for English language learners and other language minority students. *Bilingual Research Journal, 29* (2), 241-267.

Comprehensive School Reform Quality Center (2005). *CSRQ Center report on elementary school comprehensive school reform models*. Washington, DC: American Institutes for Research.

Cooper, R., Slavin, R.E., & Madden N.A. (1998). Success for All: Improving the quality of implementation of whole-school change through the use of a national reform network. *Education and Urban Society, 30*, (3), 385-408.

Cronbach, L.J., Ambron, S.R., Dornbusch, S.M., Hess, R.D., Hornik, R.C., Phillips, D.C., Walker, D.F., & Weiner, S.S. (1980). *Toward reform of program evaluation: Aims, methods, and institutional arrangements.* San Francisco, CA: Jossey-Bass.

Datnow, A., & Castellano, M. (2000). Teachers' responses to Success for All: How beliefs, experiences, and adaptations shape implementation. *American Educational Research Journal, 37* (3), 775-799.

Harris, A., Hopkins D., Youngman, M., & Wordsworth, J. (2001). The implementation and impact of Success for All in English schools. In R.E. Slavin (Ed.), *Success for All: Research and reform in elementary education*. Mahwah, NJ: Erlbaum.

Herman, R. (1999). *An educator's guide to schoolwide reform.* Arlington, VA: Educational Research Service.

Hertz-Lazarowitz, R. (2001). Success for All in Israel: A community group investigation research and evaluation model. In R.E. Slavin (Ed.), *Success for All: Research and reform in elementary education*. Mahwah, NJ: Erlbaum.

Hopfenberg, W.S., & Levin, H.M. (1993). *The Accelerated Schools resource guide*. San Francisco: Jossey-Boss.

Hopkins, D., Youngman, M., Harris, A., & Wordsworth, J. (1999). Evaluation of the initial effects and implementation of Success for All in England. *Journal of Research in Reading, 22* (3), 257-270.

Hurley, E.A., Chamberlain, A., Slavin, R.E., & Madden, N.A. (2001). Effects of Success for All on TAAS Reading: A Texas statewide evaluation. . *Phi Delta Kappan, 82* (10), 750-756.

Livingston, M., & Flaherty, J. (1997). *Effects of Success for All on reading achievement in California schools.* Los Alamitos, CA: WestEd.

Madden, N. A., Daniels, C., Chamberlain, A., Cheung, A., & Slavin, R. E. (2011). *The Reading Edge: Evaluation of a cooperative learning intervention for urban high schools*. Baltimore: Success for All Foundation.

Madden, N. A., Slavin, R. E., Logan, M., & Cheung, A. (2011). Effects of cooperative writing with embedded multimedia: A randomized experiment. *Effective Education, 3*(1), 1–9.

Madden, N.A., Slavin, R.E., Karweit, N.L., Dolan, L.J., & Wasik, B.A. (1993). Success for All: Longitudinal effects of a restructuring program for inner-city elementary schools. *American Educational Research Journal, 30,* 123-148.

Mason, B. (2005). Achievement effects of five comprehensive school reform designs implemented in Los Angeles Unified School District. (Doctoral dissertation, Pardee Rand Graduate School, 2005). Available at www.rand.org.

McLaughlin, M.W. (1990). The Rand change agent study revisited: Macro perspectives and micro realities. *Educational Researcher, 19*(9), 11-16.

Muñoz, M.A. & Dossett, D. (2004). Educating students placed at risk: Evaluating the impct of Success for All in urban settings. *Journal of Education for Students Placed at Risk, 9(3),* 261-277.

National Reading Panel. (2000) *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.

National Academy of Sciences (1998). *The prevention of reading difficulties in young children.* Washington, DC: Author.

Nunnery, J., Ross, S., Smith, L., Slavin, R., Hunter, P., & Stubbs, J. (1996, April). *An assessment of Success for All program 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.

Phenix, D., Siegel, D., Zaltsman, A., & Fruchter, N. (2004). *Virtual district, real improvement: A retrospective evaluation of the chancellor's District, 1996-2003.* New York: New York University.

Rakow, J., & Ross, S.M. (1997*). Teacher survey: Success for All, Little Rock City Schools, 1996-97.* Memphis: University of Memphis, Center for Research in Educational Policy.

Ross, S.M., Alberg, M., & McNelis, M. (1997). *Evaluation of elementary school school-wide programs: Clover Park School District year 1: 1996-97.* Memphis: University of Memphis, Center for Research in Educational Policy.

Ross, S.M., Alberg, M., McNelis, M., & Smith, L. (1998). *Evaluation of elementary school-wide programs: Clover Park School District, year 2: 1997-98*. Memphis: University of Memphis, Center for Research in Educational Policy.

Ross, S.M., Nunnery, J., & Smith, L.J. (1996). *Evaluation of Title I Reading Programs: Amphitheater Public Schools. Year 1: 1995-96.* Memphis: University of Memphis, Center for Research in Educational Policy.

Ross, S.M., Nunnery, J., & Smith, L.J. (1996). *Evaluation of Title I Reading Programs: Amphitheater Public Schools. Year 1: 1995-96.* Memphis: University of Memphis, Center for Research in Educational Policy.

Ross, S.M., Sanders, W.L., & Wright, S.P. (1998). *An analysis of Tennessee Value Added Assessment (TVAAS) performance outcomes of Roots and Wings schools from 1995-1997.* Memphis: University of Memphis.

Ross, S.M., Smith, L.J., & Casey, J.P. (1997). Preventing early school failure: Impacts of Success for All on standardized test outcomes, minority group performance, and school effectiveness. *Journal of Education for Students Placed at Risk, 2,* (1), 29-53.

Ross, S.M., Smith, L.J., & Nunnery, J.A. (1998, April). *The relationship of program implementation quality and student achievement.* Paper presented at the annual meeting of the American Educational Research Association, San Diego.

Ross, S.M., Smith, L.J., Lewis, T., & Nunnery, J. (1996). *1995-96 evaluation of Roots & Wings in Memphis City Schools.* Memphis: University of Memphis, Center for Research in Educational Policy.

Ross, S.M., Smith, L.J., Nunnery, J., & Sterbin, A. (1995). *Fall 1995 teacher survey results for the Memphis City Schools restructuring design.* Memphis, TN: University of Memphis Center for Research in Educational Policy.

Ross, S.M., Wang, L.W., Sanders, W.L., Wright, S.P., & Stringfield, S. (1999). *Two- and three-year achievement results on the Tennessee Value-Added Assessment System for restructuring schools in Memphis*. Memphis: University of Memphis, Center

for Research in Educational Policy.

Ross, S.M., Wang, L.W., Sanders, W.L., Wright, S.P., & Stringfield, S. (1999). *Two- and three-year achievement results on the Tennessee Value-Added Assessment System for restructuring schools in Memphis*. Memphis: University of Memphis, Center for Research in Educational Policy.

Rowan, B., Correnti, R., Miller, R. J., & Camburn, E. (2009). School improvement by design: Lessons from a study of comprehensive school reform designs. In B. Schnieder & G. Sykes (Eds.), *Handbook of Education Policy Research*. London: Taylor & Francis.

Russ, J., & Harris, A. (2005). Success for All: An evaluation summary. Warwick, England: University of Warwick.

Schacter, J. (1999). *Reading programs that work: A review of programs for pre-kindergarten to fourth grade*. Santa Monica, CA: Milken Family Foundation.

Slavin, R.E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research, 57*, 347-350.

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice* (2nd Ed.). Boston: Allyn & Bacon.

Slavin, R.E. (1996). Neverstreaming: Preventing learning disabilities. *Educational Leadership, 53* (5), 4-7.

Slavin, R. E. (2006). *Success for All and academically talented students*. Baltimore: Johns Hopkins University, Center for Research and Reform in Education.

Slavin, R. E., Chamberlain, A., Daniels, C., & Madden, N. A. (2009). The Reading Edge: A randomized evaluation of a middle school cooperative reading program. *Effective Education, 1*(1), 13–26.

Slavin, R.E., Cheung, A., Groff, C., and Lake, C. (2008). Effective reading programs for middle and high schools: A best evidence synthesis. *Reading Research Quarterly*, 43, 3, 290-322.

Slavin, R.E., Daniels, C., & Madden, N.A. (2005). The Success for All Middle School: Adding content to middle grades reform. *Middle School Journal*, 36 (5), 4-8.

Slavin, R. E., Karweit, N.L., & Madden, N.A. (1989) (Eds.). *Effective programs for students at risk.* Boston: Allyn & Bacon.

Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research,* 79 (4), 1391-1466.

Slavin, R. E., Lake, C., Davis, S., & Madden, N. (2011) Effective programs for struggling readers: A best evidence synthesis. *Educational Research Review, 6,* 1-26

Slavin, R. E., & Madden, N. A. (1993, April). *Multi-site replicated experiments: An application to Success for All.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta.

Slavin, R.E., & Madden, N.A. (1999). 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, 4* (4), 393-

416.

Slavin, R.E., & Madden, N.A. (2000*).* Roots & Wings: Effects of whole-school reform on student achievement. *Journal of Education for Students Placed at Risk, 5* (1 & 2), 109-136.

Slavin, R. E., Madden, N. A., Dolan, L., Wasik, B. A., Ross, S. M., & Smith, L. J. (1994). "Whenever and wherever we choose...:" The replication of Success for All. *Phi Delta Kappan, 75*(8), 639-647.

Slavin, R. E., Madden, N. A., Karweit, N. L., Dolan, L., & Wasik, B. A. (1992). *Success for All: A relentless approach to prevention and early intervention in elementary schools*. Arlington, VA: Educational Research Service.

Slavin, R. E., Madden, N. A., Karweit, N. L., Livermon, B. J., & Dolan, L. (1990). Success for All: First-year outcomes of a comprehensive plan for reforming urban education. *American Educational Research Journal, 27*, 255-278.

Slavin, R. E., Madden, N. A., & Stevens, R. J. (1989/90). Cooperative learning models for the 3 R's. *Educational Leadership, 47*(4), 22-28.

Slavin, R. E., Madden, N.A., Chambers, B., & Haxby, B. (2009). *Two million children: Success for All.* Thousand Oaks, CA: Corwin.

Slavin, R. E., Wordsworth, J., & Jones-Hill, M. (2005). *Success for All in England: Implementation and outcomes of a comprehensive literacy reform for primary schools*. Baltimore, MD: Success for all Foundation.

Smith, L. J., & Ross, S. M., & Casey, J. P. (1994). *Special education analyses for Success for All in four cities*. Memphis: University of Memphis, Center for Research in

Educational Policy.

Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative Integrated Reading and Composition: Two field experiments. *Reading Research Quarterly, 22*, 433-454.

Success for All Foundation. (2001). *Curiosity Corner teacher's manual.* Baltimore, MD: Success for All Foundation.

Success for All Foundation. (2003). *Reading Roots teacher's manual.* Baltimore, MD: Success for All Foundation.

Success for All Foundation. (2011). *Reading Wings teacher's manual (4th edition.).* Baltimore, MD: Success for All Foundation.

Success for All Foundation. (2012a). *KinderCorner (2nd edition) teacher's manual.* Baltimore, MD: Success for All Foundation.

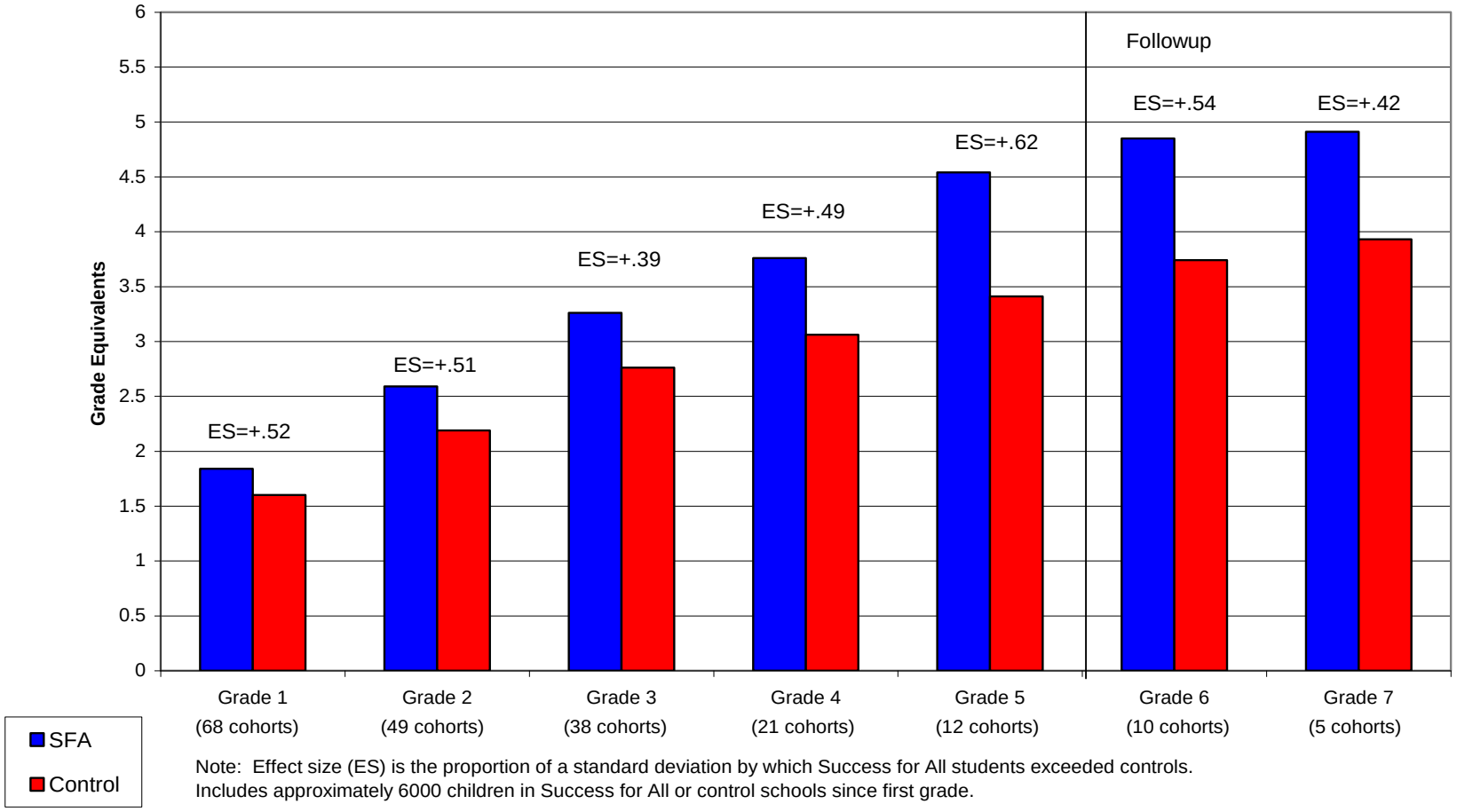
Success for All Foundation. (2012b). *Leading for Success.* Baltimore, MD: Success for All Foundation.

Success for All Foundation. (2012c). *Schoolwide Solutions: Positive plans for helping children succeed.* Baltimore, MD: Success for All Foundation.

Success for All Foundation. (2013). *Reading Edge teacher's manual.* Baltimore, MD: Success for All Foundation.

Traub, J. (1999). *Better by design? A consumer's guide to schoolwide reform*. Washington, DC: Thomas Fordham Foundation.

Figure 1
Comparison of Success for All and Control Schools in Mean Reading Grade Equivalents and Effect Sizes 1988-1999



Note: Effect size (ES) is the proportion of a standard deviation by which Success for All students exceeded controls. Includes approximately 6000 children in Success for All or control schools since first grade.

Figure 2
Houston Independent School District
1996 First Grade Effect Sizes by Implementation Level-English

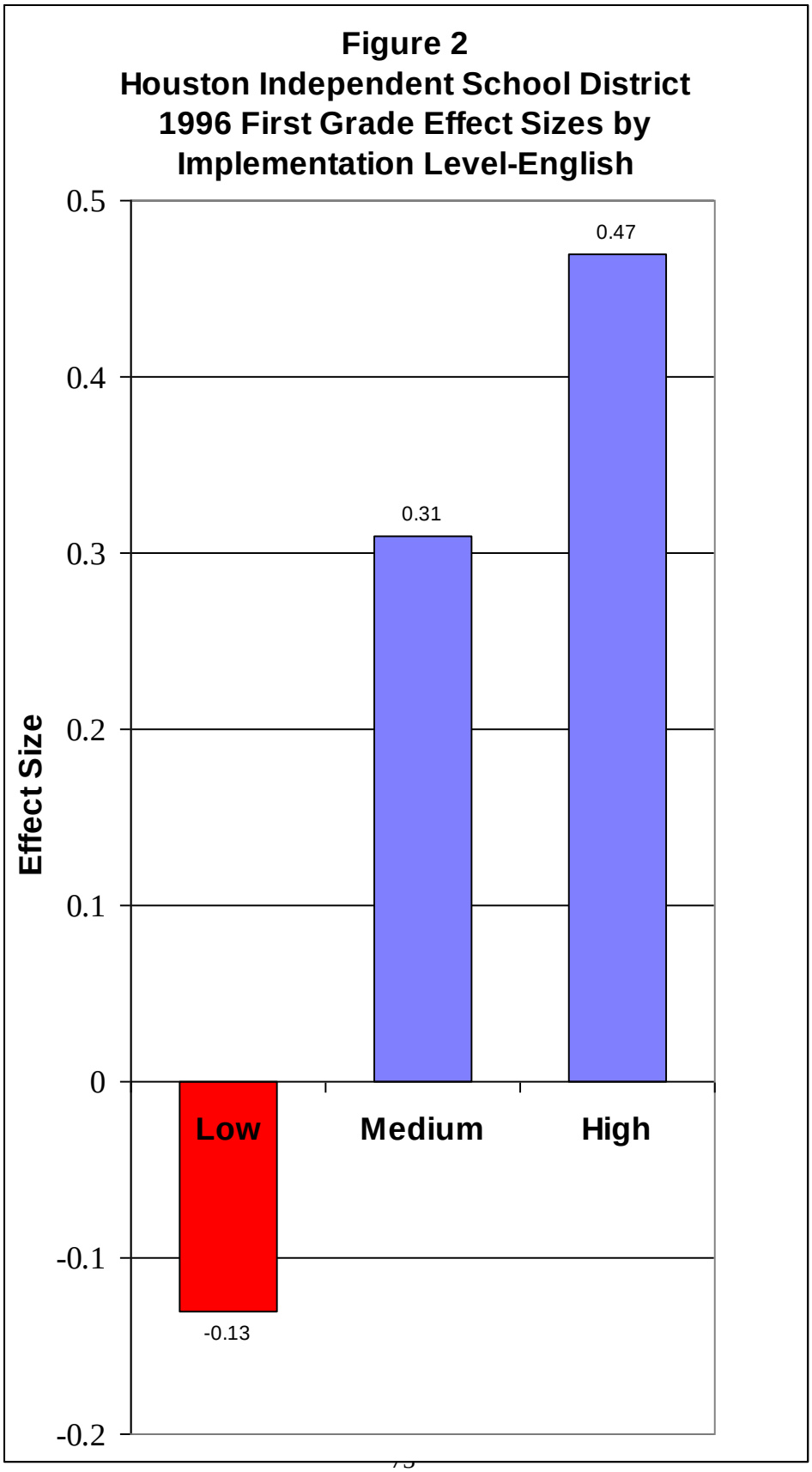


Figure 3
Houston Independent School District
1996 First Grade Effect Sizes by
Implementation Category-Spanish

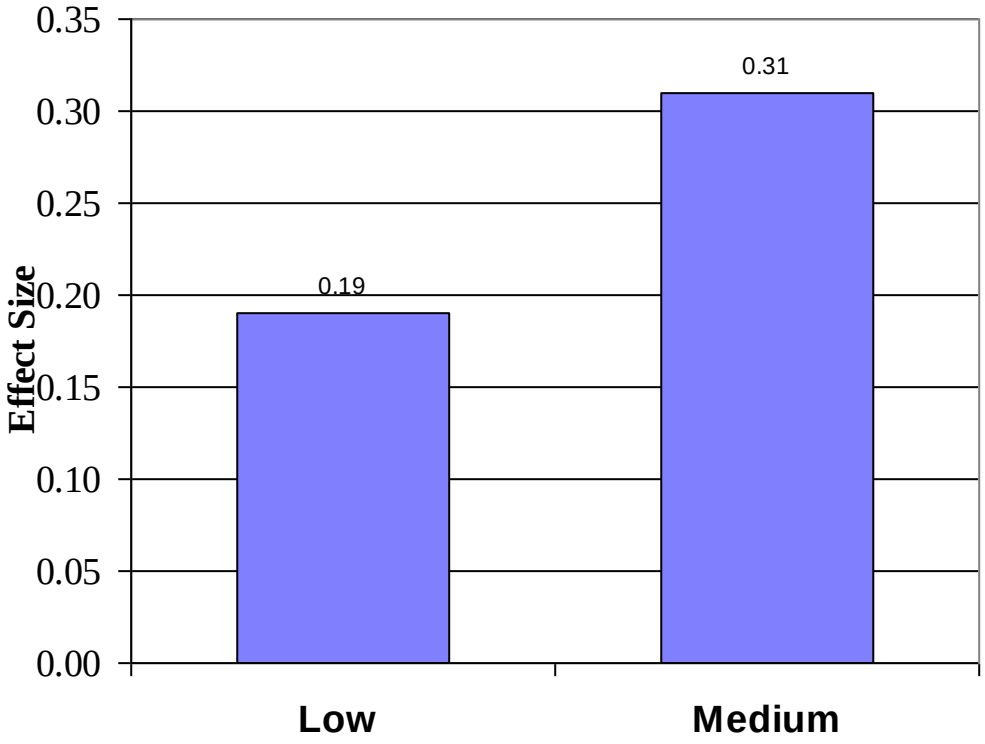
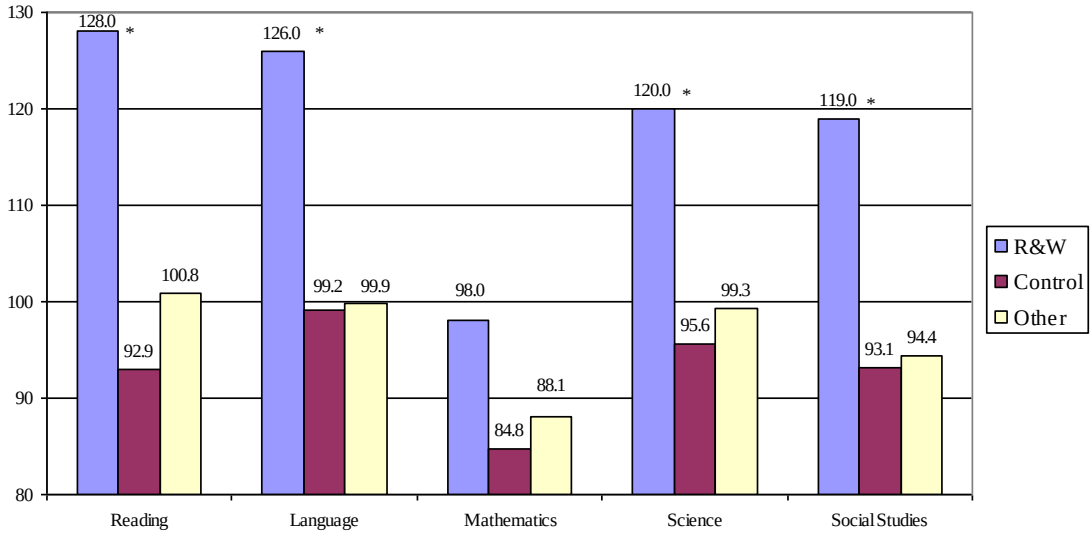


Figure 4
Percent of Expected Gain on TVAAS for Roots & Wings, Control, and Other Memphis Schools, Grades 2-5, 1997



* Roots & Wings scores significantly higher than those of controls or others (p<.05).
Tennessee Value-Added Assessment System (TVAAS) scores analyzed by William Sanders for the Tennessee State Department of Education. S
Ross, Sanders, & Wright, 1998.

Figure 5
Gains From Preimplementation Year to 1998
Success for All vs. Texas Means
Texas Assessment of Academic Success
Reading, Grades 3, 4, 5

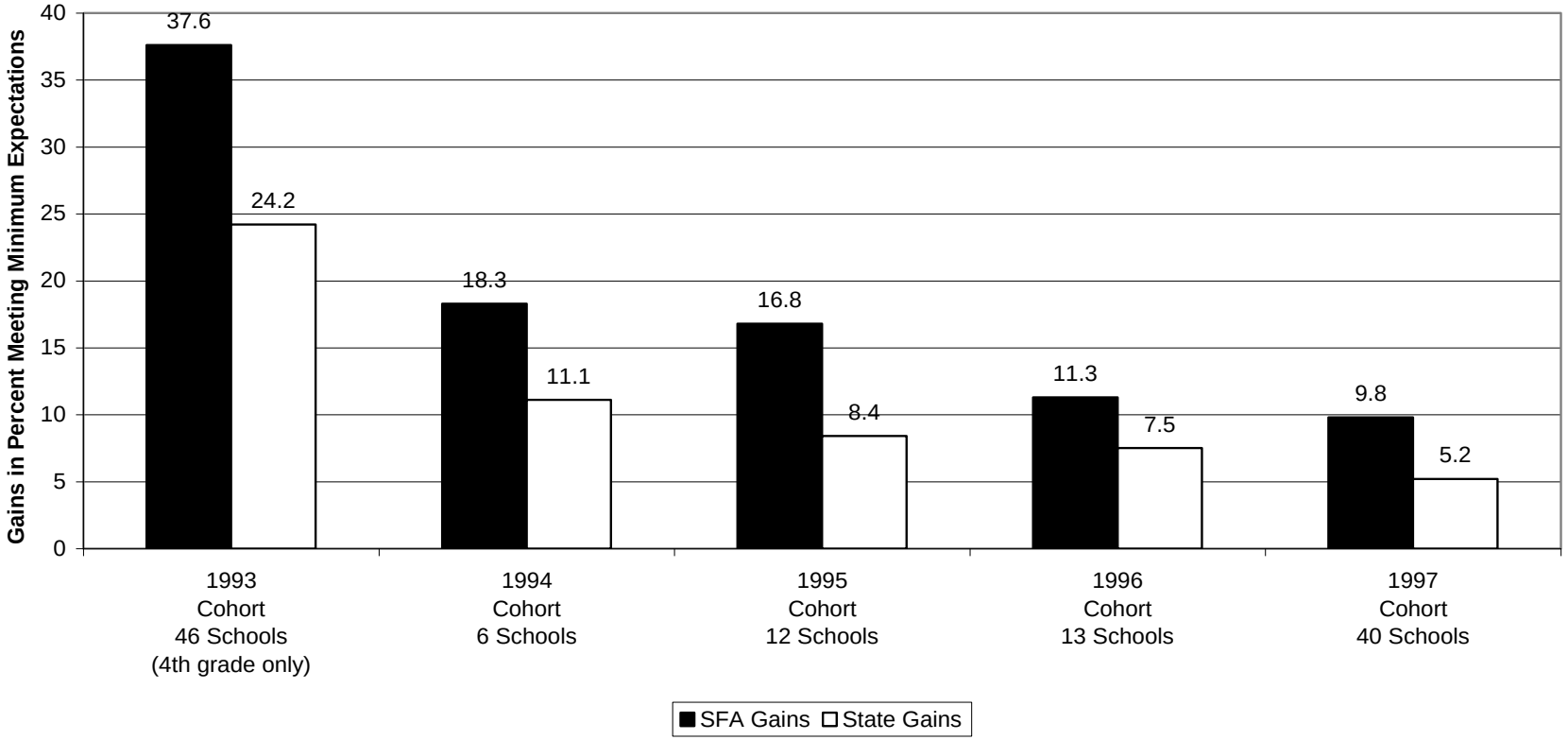


Figure 6
Effect Sizes Comparing Success for All and Control Schools According to Implementation Year