

Social Programs That Work Review

Evidence Summary for Success for All for Grades K-2

HIGHLIGHTS:

- **PROGRAM:** A school-wide reform program, primarily for high-poverty elementary schools, with a strong focus on reading instruction.
- **EVALUATION METHODS:** Two large, well-conducted randomized controlled trials (RCTs) with a combined sample of 78 high-poverty elementary schools.
- **KEY FINDINGS:** The first RCT found that the program increased second-grade reading achievement in Success for All schools by 25-30% of a grade-level, three years after random assignment. The second RCT, however, did not find significant positive effects on most measures of reading achievement, including comprehension (the most important, final measure).
- **OTHER:** This evidence of effectiveness applies to the program as implemented in grades K-2 (as opposed to later elementary school). The reasons for the discrepant findings between the two RCTs are unclear. We believe that an additional, well-conducted RCT showing positive effects on reading comprehension is needed to have confidence that the program meaningfully boosts student reading ability when implemented in typical high-poverty school settings.

I. Evidence rating: **SUGGESTIVE TIER**

The standard for Suggestive Tier is:

Programs that have been evaluated in one or more well-conducted RCTs (or studies that closely approximate random assignment) and found to produce sizable positive effects, but whose evidence is limited by only short-term follow-up, effects that fall short of statistical significance, or other factors. Such evidence suggests the program may be an especially strong candidate for further research, but does not yet provide confidence that the program would produce important effects if implemented in new settings.

II. Description of the Program:

Success for All is a comprehensive school-wide reform program, primarily for high-poverty elementary schools, with a strong emphasis on early detection and prevention of reading problems before they

become serious. Key program elements include: (i) daily 90-minute reading classes, each of which is formed by grouping together students of various ages who read at the same performance level; (ii) a K-1 reading curriculum that focuses on language development (e.g., reading stories to students and having them re-tell), teaching students the distinct sounds that make up words (i.e. phonemic awareness), blending sounds to form words, and developing reading fluency; (iii) daily one-on-one tutoring (in addition to the classes) for students needing extra help with reading; and (iv) cooperative learning activities (in which students work together in teams or pairs) starting in the grade 2 reading classes.

The program's cost, as measured in the first RCT, was approximately \$600 per student, over the full three-year period (in 2017 dollars). In the second RCT, the cost of the program over the full three-year period was approximately \$900 per student (in 2017 dollars). These cost estimates includes materials and training to implement the program in grades K-5. Schools may incur additional costs of reallocating staff from other functions (e.g., to provide a higher ratio of tutors).

[Click here for Success for All's website.](#)

III. Evidence of Effectiveness:

This program was evaluated in two well-conducted RCTs. The following summarizes the program's effects on the main outcomes measured in each study. Importantly, both trials evaluated the program as implemented in grades K-2 but not grades 3-5; thus, the findings apply only to the K-2 elements of the program.

STUDY 1

Study Design: Large, multi-site randomized controlled trial evaluating Success For All in a sample of 41 high-poverty elementary schools across 11 states, during 2001-2006.

Prior to random assignment, at least 80% of the schools' teachers had voted in favor of adopting Success for All (a step that the program provider typically requires before working with a school), and the schools had agreed to allow data collection over the course of the study. The 41 schools were randomly assigned either to a group that implemented Success for All in grades K-2 or a control group that did not (most implemented the program in grades 3-5 instead). The schools contained a total of 2,694 entering kindergarten students administered a pretest at the start of the study. The student population in these 41 schools was 56% African-American and 10% Hispanic, and 72% of students were low-income (i.e., eligible for federally subsidized lunches).

Approximately three years after random assignment, the study assessed reading outcomes for all second-grade students in the sample schools. Sixty-nine percent of these students had been exposed to Success for All, or the control condition, for all three years of the study (i.e., in grades K-2); the other 31% had enrolled in the Success for All or control schools during the study, and so had received partial exposure.

Effects of Success for All on school-wide second-grade reading outcomes, three years after random assignment (versus the control schools):

These are the effects on all outcomes that the study measured at the three-year follow-up. All effects shown are statistically significant at the 0.05 level unless stated otherwise.

- On average, second graders at Success for All schools –
 - › Scored higher in passage comprehension than approximately 58% of their counterparts at control group schools (this equates to a standardized effect size of 0.21).
 - › Scored higher in word identification skills than approximately 60% of their counterparts at control group schools (this equates to a standardized effect size of 0.24); and
 - › Scored higher in word attack skills than approximately 64% of their counterparts at control group schools (this equates to a standardized effect size of 0.36).
- To express these effects as grade level equivalents: On average, second graders at Success for All schools score approximately 25-30% of a grade level higher in reading ability than their counterparts at the control schools.¹
- The program's effects generally grew in size from the first to the third year of the study (the earlier, smaller effects were mostly not statistically significant) –
 - › Passage comprehension: From an effect size of -0.10 in year 1, to 0.12 in year 2, to 0.21 in year 3.
 - › Word identification skills: From 0.09 to 0.19 to 0.24.
 - › Word attack skills: From 0.32 to 0.29 to 0.36.

(The earlier, smaller effects were mostly not statistically significant.)

Summary of Study Quality:

- This was a large, multi-site study evaluating Success for All as it is typically implemented in high-poverty elementary schools, thus providing evidence about the program's effectiveness in real-world public school settings.
- The study had a reasonably long-term follow-up, and low-to-moderate attrition: Three years after random assignment, reading test scores were obtained for students in 85% of the sample schools – i.e., 35 of the original 41. (Of the 6 schools lost at follow-up, 5 closed due to insufficient enrollment and 1 dropped the Success for All model due to local political problems and refused

¹ Specifically, the average annual gain in reading for U.S. students between the end of first grade and the end of second grade on seven nationally normed tests is 0.97, expressed as a standardized effect size (see Bloom, Hill, and Lipsey 2008, referenced at the end of this summary). The reading improvement in Success for All schools compared to the control group, summarized above under the first main bullet, represents a gain of about 25-30% of this amount.

to participate in data collection.) The number of schools lost in the Success for All versus control group were the same (3 each).

- The 35 Success for All and control schools remaining in the sample at the three-year follow-up were highly similar in their observable pre-program characteristics (e.g., average receptive vocabulary score, school enrollment, demographics).
- The study measured outcomes for all students in the second grade at the sample schools, regardless of the amount of exposure they had had to the program (i.e., the study used an “intention-to-treat” analysis).
- The study measured reading outcomes using tests whose reliability and validity are well-established (namely, the Woodcock Reading Mastery Test-Revised, subtests on word attack, word identification, and passage comprehension).
- These tests were administered by trained testers who were blind as to whether students attended Success for All or control schools.
- The study’s statistical analysis accounted for the fact that schools, rather than individual students, were randomly assigned to the Success for All versus control group.
- The results for the subsample of individual students who remained in the study for all three years were approximately the same as the school-wide results summarized above, providing a degree of corroboration for the school-wide results.

STUDY 2

Study Design: Large, multi-site randomized controlled trial evaluating Success For All in a sample of 37 high-poverty elementary schools across four states, during 2011-2014.

This was a well-conducted RCT of Success for All as it was implemented in grades K-2 under a scale-up grant from the Department of Education’s Investing in Innovation Fund. Thirty-seven high-poverty elementary schools in four states were randomly assigned to either a group that implemented Success for All school-wide or a control group that did not. The study tracked outcomes for the cohort of students in kindergarten at the start of the study (n = 2,956 students at baseline). The year 3 follow-up’s primary analysis examined effects for sample students who had at least one valid test score at the kindergarten, first grade, and second grade follow-ups.

Key Findings:

The study’s key findings at the three-year post-random assignment follow-up (i.e., end of second grade) are summarized in the table below. For reference, the table also shows the three-year findings from Study 1.

As shown in the table, in Study 2, Success for All was found to significantly increase students’ word attack skills (effect size of 0.15 standard deviations, $p < .05$), although the effect was much smaller than the 0.35 standard deviation increase in these skills found in year 2. Success for All had no

significant effects on letter-word identification, fluency, or comprehension (the non-significant effects favored the treatment group but were small in magnitude – e.g., 0.03 for comprehension, the most important, final measure of reading ability). These findings were consistent across a range of demographic and socio-economic subgroups.

The effects found in Study 2 are considerably weaker than those found in Study 1, where there was continual growth in the program’s effect on word identification and word attack skills over the three-year study period, and a significant positive effect on comprehension in year 3 (effect size of 0.21, $p < 0.05$). The reasons for the discrepant findings are unclear. Challenges implementing Success for All in Study 2 – in particular, incomplete implementation of the program’s tutoring intervention in year 3 due to funding cutbacks related to the national economic recession – may partially explain its weaker effect findings. However, Study 1 also experienced some challenges in program implementation, including inadequate delivery of tutoring at most schools in the treatment group.

Success for All’s End-of Second Grade Effect Sizes in Standard Deviations (SD)

Outcomes (Year 3)	Study 2	Study 1
Letter-Word ID	0.07 SD	Not measured (letter) 0.22 SD* (word)
Word Attack	0.15 SD*	0.33 SD*
Reading Efficiency	0.07 SD	Not measured
Passage Comprehension	0.03 SD	0.21 SD*

*statistically-significant at the 0.05 level, in a two-tailed test

Summary of Study Quality:

The study had a number of strengths: (i) Success for All and control schools were highly similar in their observable pre-program characteristics; (ii) all schools were retained in the sample through the three-year follow-up (i.e. there was no school attrition); (iii) the study measured outcomes for all students in the second grade at the sample schools, regardless of the amount of exposure they had had to the program (i.e., the study used an “intention-to-treat” analysis); (iv) the study’s statistical analysis accounted for the fact that schools, rather than individual students, were randomly assigned to Success for All or the control group; and (v) the study used well-established standardized tests to measure outcomes.

A study limitation was high overall student attrition at the year 3 follow-up (44%). However, it is reassuring that (i) the student attrition rates for the Success for All versus control group were almost identical (within half of a percentage point of each other), and (ii) the study report showed that the Success for All and control members remaining in the sample after attrition were still highly similar in their observable pre-program characteristics (demographics, as well early language and literacy skills). Both of these items suggest that attrition did not seriously undermine the equivalence of the Success for All and control groups.

OTHER STUDIES

One other RCT of Success for All has been conducted; it took place in the United Kingdom and reported findings in 2017. We do not summarize its results here because the study only measured effects on reading skills at the end of kindergarten. These effects were modest in size and not statistically significant.

IV. References:

Borman, Geoffrey D., Robert E. Slavin, Alan Cheung, Anne Chamberlain, Nancy Madden, and Bette Chambers, "Success for All: First-Year Results from the National Randomized Field Trial," *Educational Evaluation and Policy Analysis*, 2005, vol. 27, no. 1, pp. 1-22. [Click here for a link to this study.](#)

Borman, Geoffrey D., Robert E. Slavin, Alan Cheung, Anne Chamberlain, Nancy Madden, and Bette Chambers, "The National Randomized Field Trial of Success for All: Second-Year Outcomes," *American Educational Research Journal*, Winter 2005, vol. 42, no. 4, pp. 673-696.

Borman, Geoffrey D., Robert E. Slavin, Alan Cheung, Anne Chamberlain, Nancy Madden, and Bette Chambers, "Final Reading Outcomes of the National Randomized Field Trial of Success for All," *American Educational Research Journal*, September 2007, vol. 44, no. 3, pp. 701-731.

Bloom, Howard S., Carolyn Hill, Alison Rebeck Black, and Mark Lipsey, *Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions*, MDRC Working Paper on Research Methodology, October 2008.

Miller, Sarah, Andy Biggart, Seaneen Sloan, and Liam O'Hare, *Success for All – Evaluation Report and Executive Summary*, Education Endowment Foundation, 2017.

Quint, Janet, Pei Zhu, Rekha Balu, Shelley Rappaport, and Micah DeLaurentis, *Scaling Up the Success for All Model of School Reform: Final Report from the Investing in Innovation (i3) Evaluation*, MDRC, 2015.