What Works Clearinghouse

Success for All

Program description[1]

Success for All (SFA) is a comprehensive reading, writing, and oral language development program for students in pre-K through eighth grade. Its underlying premise is that all children, including those with limited English proficiency, can and should be reading in English at grade level by the end of third grade. (*SFA* can impact Spanish literacy as well, though these outcomes fall outside the scope of this report.) Initial reading instruction is delivered in 90-minute daily blocks to students grouped by reading level, across classes and grades. Certified teachers provide daily tutoring to those students who are having difficulty reading. In addition, Family Support Teams and full-time *SFA* facilitators train teachers, oversee student assessments, encourage parental involvement, work to decrease absenteeism, and assist with decisions about group placement and tutoring.

Research

One study of *SFA* met the What Works Clearinghouse (WWC) evidence standards with reservations.[2] The study included 324 English language learners from kindergarten to first grade who attended elementary schools in the District of Columbia, New York, Arizona, California, and Illinois.[3]

The WWC considers the extent of evidence for *SFA* to be small for reading achievement. No studies that met WWC evidence standards with or without reservations addressed mathematics achievement or English language development.

Effectiveness

Success for All was found to have potentially positive effects on reading achievement.

	Reading achievement	*Mathematics achievement*	*English language development*
Rating of effectiveness	Potentially positive	na	na
Improvement index[4]		na	na
	Average: +11 percentile points		
	Range: +5 to +17 percentile points		

na = not applicable

1. The descriptive information for this program was obtained from publicly available sources: the program's website (www.successforall.org) and the research literature (Chambers et al., 2004). The WWC requests that developers review the program description sections for accuracy from their perspective. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review.
2. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.
3. Subsample data for Spanish-dominant English lanugage learner (ELL) students were obtained from the study author.
4. These numbers show the average and range of improvement indices for all findings across the studies.

Developer and contact

Developed by Dr. Robert Slavin, *SFA* is distributed by Success for All Foundation, Inc. Address: 200 W. Towsontown Boulevard, Baltimore, Maryland 21204-5200. Email: sfainfo@successforall.org. Web: www.successforall.org. Telephone: (800) 548-4998 ext. 2300.

Scope of use

SFA is used by schools in 48 states, Guam, and the Virgin Islands. According to the Success for All Foundation, more than 1,300 schools in more than 500 districts use the *SFA* program. England, Israel, Canada, Mexico, and Australia have implemented adapted versions of *Success for All.* Information on the scope of use specifically for English language learners is not available.

Teaching

Teachers at each grade level begin the instructional period by reading literature to students and engaging them in story discussion to enhance comprehension, listening and speaking vocabulary, and knowledge of story structure. In kindergarten and first grade, teachers emphasize the development of reading and language skills by using storybooks and instruction on phonemic awareness, auditory discrimination, and sound blending.

For most of the day, students in grades 1 through 6 are assigned to heterogeneous, age-grouped classes comprised of approximately 25 students. However, the main component of the *SFA* program occurs during a regular 90-minute reading period, when they are regrouped into reading classes of 15–20 students who all perform at the same reading level. Regrouping allows teachers to teach without having to break the class into reading groups. The study reviewed (Chambers et al., 2004) includes *Reading Reels*, which are video materials developed to enhance the *SFA* program. During the lessons, teachers show their students 30-second to three-minute video skits demonstrating elements of beginning reading. For six months, *Reading Reels* were used daily to supplement *SFA* instruction.

Cost

The cost for the *SFA* program for a school of 500 children is approximately $80,000 in the first year of implementation; the price decreases in subsequent years. The cost is $40,000 in the second year of implementation and $30,000 in the third year. The cost of the intervention includes materials and training for staff.

Twenty studies reviewed by the WWC investigated the effects of *SFA* for English language learners who are taught to read in English. One of the studies (Chambers, Slavin, Madden, Cheung, & Gifford, 2004) was a quasi-experimental design that met WWC evidence standards with reservations. The remaining 19 studies did not meet WWC evidence screens.

Met evidence standards with reservations

Chambers et al. (2004) was a quasi-experimental design that studied 455 kindergarten and first grade Hispanic students from four states and the District of Columbia. Of these students, 324 were English language learners. This report is based on data pertaining to the English language learners only. The study examined the effects of using the program with embedded video (*Reading Reels*) on reading achievement. Four of the eight schools involved in the study (four experimental and four comparison schools) were included in a national randomized evaluation of *SFA* (Borman et al., 2003).[5]

Extent of evidence

The WWC categorizes the extent of evidence in each domain as small or moderate to large (see the What Works Clearinghouse Extent of Evidence Categorization Scheme). The extent of evidence takes into account the number of studies and the

5. Forty-one high-poverty schools participated in a national study of *Success for All.*

Research *(continued)*

total sample size across the studies that met WWC evidence standards with or without reservations.[6]

The WWC considers the extent of evidence for *Success for All* to be small for reading achievement. No studies that met WWC evidence standards with or without reservations addressed mathematics achievement or English language development.

Effectiveness

Findings

The WWC review of interventions for English language learners addresses student outcomes in three domains: reading achievement, mathematics achievement, and English language development. The study reviewed assessed outcomes in reading achievement.

Reading achievement. Chambers et al. (2004) reported statistically significant findings based on three outcome measures of reading achievement (Woodcock Reading Mastery Test: Word Identification, Word Attack, and Passage Comprehension), but none of these outcomes were statistically significant according to the WWC analysis. The average effect size across the three student outcomes was, however, large enough to be considered substantively important according to WWC criteria (that is, at least 0.25).

Rating of effectiveness

The WWC rates the effects of an intervention in a given outcome domain as: positive, potentially positive, mixed, no discernible effects, potentially negative, or negative. The rating of effectiveness takes into account four factors: the quality of the research design, the statistical significance of the findings, the size of the difference between participants in the intervention and the comparison conditions, and the consistency in findings across studies (see the WWC Intervention Rating Scheme).[7]

The WWC found *Success for All* to have potentially positive effects on reading achievement for English language learners

Improvement index

The WWC computes an improvement index for each individual finding. In addition, within each outcome domain, the WWC computes an average improvement index for each study as well as an average improvement index across studies (see Technical Details of WWC-Conducted Computations). The improvement index represents the difference between the percentile rank of the average student in the intervention condition versus the percentile rank of the average student in the comparison condition. Unlike the rating of effectiveness, the improvement index is based entirely on the size of the effect, regardless of the statistical significance of the effect, the study design, or the analyses. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.

The average improvement index for reading achievement is +11 percentile points, with a range of +5 to +17 percentile points across findings in the one study.

Summary

The WWC reviewed 20 studies of the effectiveness of *SFA* for English language learners learning to read in English. One study met WWC evidence standards with reservations; the remaining studies did not meet WWC evidence screens. Based on this

---

6. The Extent of Evidence Categorization was developed to tell readers how much evidence was used to determine the intervention rating, focusing on the number and size of studies. Additional factors associated with a related concept, external validity, such as the students' demographics and the types of settings in which studies took place, are not taken into account for the categorization.

7. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate the statistical significance. In the case of *Success for All,* corrections for clustering and multiple comparisons were needed.

one study, the WWC found potentially positive effects in reading achievement. The evidence presented in this report may change as new research emerges.

References

Met WWC evidence standards with reservations

Chambers, B., Slavin, R. E., Madden, N. A., Cheung, A., & Gifford, R. (2004). *Effects of Success for All with embedded video on the beginning reading achievement of Hispanic children.* Baltimore, MD: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.
  Additional source:
  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,* 125–230.

Did not meet WWC evidence screens

Chambers, B., Cheung, A., Gifford, R., Madden, N., & Slavin, R. E. (2004). *Achievement effects of embedded multimedia in a Success For All reading program.* Baltimore, MD: Success for All Foundation.[8]
Dianda, M. R., & Flaherty, J. F. (1995). *Report on work station uses: Effects of Success for All on the reading achievement of first graders in California bilingual programs.* Los Alamitos, CA. (ERIC Document Reproduction Service No. ED394327)[9]
Hurley, E. A., Chamberlain, A., Slavin, R. E., & Madden, N. A. (2001). Effects of Success for All on TAAS reading scores—A Texas statewide evaluation. *Phi Delta Kappan, 82*(10), 750–756.[10]

Livingston, M., & Flaherty, J. (1997). *Effects of Success for All on reading achievement in California schools.* Los Alamitos, CA: WestEd.[11]
Nunnery, J. A. (1995). An assessment of Success for All program component effects on the reading achievement of at-risk first-grade students. *Dissertation Abstracts International, 57*(01), 155A. (UMI No. 9615378)[12]
Nunnery, J. A., Slavin, R. E., Madden, N. A., Ross, S. M., Smith, L. J., Hunter, P., et al. (1997, March). *Effects of full and partial implementations of Success for All on student reading achievement in English and Spanish.* Paper presented at the meeting of the American Educational Research Association, Chicago, IL.[12]
Ross, S. M., & Casey, J. (1998). *Longitudinal study of student literacy achievement in different Title I school-wide programs in Fort Wayne community schools—Year 2: First grade results.* Memphis, TN: University of Memphis, Center for Research in Educational Policy.[10]
Ross, S. M., Nunnery, J. A., & Smith, L. J. (1996). *Evaluation of Title I reading programs: Amphitheater public schools—Year 1: 1995–1996.* Memphis, TN: University of Memphis, Center for Research in Educational Policy.[10]
Ross, S. M., Smith, L. J., & Nunnery, J. A. (1998, April). *Title I as a catalyst for school improvement: Impact of alternative school-wide models on the reading achievement of students*

8. The effects of two interventions were studied: the WWC could not examine the results for the intervention that is relevant to this review.
9. Confound: the analysis could not separate the effects of the intervention from the effects of the school.
10. The sample is not appropriate to this review: data were not disaggregated so that the WWC could not examine the results for the sample that is relevant to this review.
11. The language of instruction differed between groups.
12. The outcome measures are not relevant to this review.

References *(continued)*

at risk. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.[10]

Slavin, R. E., Leighton, M., & Yampolsky, R. (1990). *Success For All: Effects on the achievement of limited English proficient children* (Report No. 5). Baltimore, MD: The Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students. (ERIC Document Reproduction Service No. ED331585)[13]

Slavin, R. E., & Madden, N. A. (1998). *Success for All / Éxito Para Todos—Effects on the reading achievement of students acquiring English* (Report No. 19). Baltimore, MD: Center for Research on the Education of Students Placed at Risk. (ERIC Document Reproduction Service No. ED423327)[14]

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.[15]

Slavin, R. E., Madden, N. A., Dolan, L., Wasik, B. A., Ross, S. M., & Smith, L. J. (1994, April). *Success for All: Longitudinal effects of systemic school-by-school reform in seven districts.* Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.[16]

Slavin, R. E., & Yampolsky, R. (1991). *Effects of Success for All on students with limited English proficiency: A three year evaluation* (Draft). Baltimore, MD: Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students.[13]

Slavin, R. E., & Yampolsky, R. (1991). *Success for All: Effects on language minority students.* Baltimore, MD: The Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students.[13]

Slavin, R. E., & Yampolsky, R. (1992). *Success for All: Effects on students with limited English proficiency: A three year evaluation* (Report No. 29). Baltimore, MD: Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students. (ERIC Document Reproduction Service No. ED346199)[13]

Urdegar, S. M. (1998). *Evaluation of the Success for All program 1997–98.* Miami, FL: Miami-Dade Public Schools, Office of Educational Evaluation.[17]

Urdegar, S. M. (2000). *Evaluation of the Success for All program 1998–99.* Miami, FL: Miami-Dade County Public Schools, Office of Evaluation and Research.[17]

Wang, L. W., & Ross, S. M. (2003). *Comparisons between elementary school programs on reading performance: Albuquerque Public Schools.* Memphis, TN: The University of Memphis, Center for Research in Educational Policy.[10]

For more information about specific studies and WWC calculations, please see the WWC *Success for All Technical Appendices*.

13. Confound: there was only one school in each study condition, so the analysis could not separate the effects of the intervention from the effects of the school.
14. Incomparable groups: there were differences in the amount of native language used.
15. For studies in which outcomes of the intervention appropriate to this review were examined, two had confounds: there was one school in one of the study conditions (Arizona) or one school in each study condition (Francis Scott Key), so the analysis could not separate the effects of the school. Complete data were not reported in one study (El Vista), so the WWC could not calculate effect sizes.
16. For studies in which outcomes of the intervention appropriate to this review were examined, the samples are not appropriate to the review: data were not disaggregated, so the WWC could not examine the results of the sample that is relevant to this review.
17. Complete data were not reported: the WWC could not compute effect sizes.

Appendix

Appendix A1   Study characteristics: Chambers, Slavin, Madden, Cheung, & Gifford, 2004 (quasi-experimental design)

Characteristic	Description
Study citation	Chambers, B., Slavin, R. E., Madden, N. A., Cheung, A., & Gifford, R. (2004). *Effects of Success for All with embedded video on the beginning reading achievement of Hispanic children.* Baltimore, MD: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.
Participants	The study included 324 Hispanic students in kindergarten and the first grade.[1] Three-quarters of the students in schools involved in the study qualified for free lunch. Each of the four intervention schools were matched with similar comparison schools.
Setting	Eight schools located in high-poverty districts in the District of Columbia, New York, Arizona, California, and Illinois.
Intervention	For six months, *Reading Reels,* a series of video segments that each last up to three minutes, were shown during lessons for a maximum of six minutes. Skits and animation in the video segments demonstrate components of beginning reading and were used in conjunction with *Success for All.* The *SFA* program included a reading curriculum, tutoring, quarterly assessments, family support teams for students' parents, school personnel facilitator, and intervention teacher training. *Reading Reels* used video to teach a concept, such as defining a word and sounding it out. Students were then asked to demonstrate the concept, typically by imitating the puppets in the video sounding out a word, both individually and with partners.
Comparison	Teachers in comparison schools taught their students with the school's existing curricula.
Primary outcomes and measurement	Although four subtests of the Woodcock Reading Mastery Test were administered, only the three that require actual reading are included in this report: Word Identification, Word Attack, and Passage Comprehension. Letter Identification was excluded because it is not a measure of reading achievement. (See Appendix A2 for more detailed descriptions of outcome measures.)
Teacher training	During their first year providing *SFA* instruction, teachers participate in training for three days during the summer, and two to four additional in-service days during the school year. Throughout each year, additional in-service presentations covering classroom management, instructional pace, and cooperative learning are provided by school facilitators and other *SFA* staff. Facilitators organize information sessions in a manner that allows teachers to share problems and solutions, suggest changes, and discuss individual children. Twice a year, trainers conduct teacher observations. After the first year, training is reinforced by regular in-services, an annual *SFA* conference, and implementation checks for the facilitators and trainers. The staff development model used in the program emphasizes relative brief initial training with extensive classroom follow-up, coaching, and group discussion. Teachers also received training about showing the videos in a sample lesson.

1.   Subsample data for Spanish-dominant ELL students were obtained from the study author.

Appendix A2   Outcome measures in the reading achievement domain

Outcome measure	Description
Woodcock Reading Mastery Test (WRMT): Word Identification	This subtest measures basic word reading skills. The standardized test requires the child to read aloud isolated real words that range in frequency and difficulty (as cited in Chambers, Slavin, Madden, Cheung, & Gifford, 2004).
WRMT: Word Attack	This subtest measures phonemic decoding skills by asking students to read pseudowords. For this standardized test, students are made aware that the words are not real (as cited in Chambers, Slavin, Madden, Cheung, & Gifford, 2004).
WRMT: Passage Comprehension	This subtest measures comprehension by having students fill in missing words in a short paragraph. This is a standardized test (as cited in Chambers, Slavin, Madden, Cheung, & Gifford, 2004).

| Outcome measure | Study sample | Sample size (schools/ students)[3] | Authors' findings from the study | | WWC calculations | | | |
| | | | Mean outcome (standard deviation[2]) | | | | | |
			Success for All group	Comparison group	Mean difference[4] (SFA – comparison)	Effect size[5]	Statistical significance[6] (at $\alpha = 0.05$)	Improvement index[7]
Chambers, Slavin, Madden, Cheung, and Gifford, 2004 (quasi-experimental design)[8]								
WMRT: Word Identification	Kindergarten & first grade	8/324	414.29 (38.92)	404.83 (29.14)	9.46	0.26	ns	+10
WMRT: Word Attack	Kindergarten & first grade	8/324	476.59 (21.91)	467.23 (16.38)	9.36	0.45	ns	+17
WMRT: Passage Comprehension	Kindergarten & first grade	8/324	446.55 (23.03)	443.81 (20.65)	2.74	0.12	ns	+5
Domain average[9] for reading achievement (Chambers et al., 2004)						0.28	ns	+11

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Subsample data for Spanish-dominant ELL students were obtained from the study author.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition versus the percentile ranking of the average student in the comparison condition. The improvement index can take on values between −50 and +50, with positive numbers denoting results favorable to the intervention group.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Chambers et al. (2004), corrections for clustering and multiple comparisons were needed. Therefore, the significance levels differ from those reported in the original study.
9. This row provides the study average, which in this instance is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A4   *Success for All* rating for the reading achievement domain

The WWC rates an intervention's effects in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of reading achievement, the WWC rated *Success for All* as potentially positive. It did not meet the criteria for positive effects because there was only one study and it did not meet WWC evidence standards for a strong design. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, negative effects) were not considered because *Success for All* was assigned the highest applicable rating.

Rating received

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

  Met. *Success for All* met this criterion because the study reviewed had substantively important positive effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

  Met. *Success for All* met this criterion because the one study reviewed did not have statistically significant or substantively important negative findings.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. *Success for All* did not meet this criterion because only one study was reviewed.

AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

  Met. *Success for All* met this criterion because the study reviewed did not have statistically significant or substantively important negative findings.

---

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.

Appendix A5   Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence[1]
		Schools	Students	
Reading achievement	1	8	324	Small
Mathematics achievement	0	0	0	na
English language development	0	0	0	na

na = not applicable/not studied

1. A rating of "moderate to large" requires at least two studies and two schools across studies in one domain, and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is "small."