

What Works Clearinghouse



Reading Edge

Program Description¹

Reading Edge is a middle school literacy program that emphasizes cooperative learning, goal setting, feedback, classroom management techniques, and the use of metacognitive strategy, whereby students assess their own skills and learn to apply new ones. The program is a component of the *Success for All*² (SFA)[®] whole-school reform model and provides eight levels of instruction, from beginning through eighth-grade reading levels. Students are grouped into classes based on ability, and whole-class reading instruction is delivered in daily 60-minute blocks. Instruction at the early levels uses fiction, nonfiction, and simple scripts to help students develop basic decoding skills, reading fluency, vocabulary, and comprehension. At reading level 3 and higher, students focus on developing comprehension strategies using both narrative and expository texts. All levels focus on building background knowledge and developing study skills. Although the program is often implemented in the context of the SFA[®] whole-school reform, this report focuses on *Reading Edge* as a stand-alone program in grades 4 and higher.³

Report Contents	
Overview	p. 1
Program Information	p. 2
Research Summary	p. 3
Effectiveness Summary	p. 4
References	p. 5
Research Details for Each Study	p. 9
Outcome Measures for Each Domain	p. 11
Findings Included in the Rating for Each Outcome Domain	p. 12
Supplemental Findings for Each Outcome Domain	p. 13
Endnotes	p. 14
Rating Criteria	p. 15
Glossary of Terms	p. 16

Research⁴

One study of *Reading Edge* that falls within the scope of the Adolescent Literacy review protocol meets What Works Clearinghouse (WWC) evidence standards without reservations, and no studies meet WWC evidence standards with reservations. The one study included 405 sixth-grade students who attended one middle school in Florida and one middle school in West Virginia.⁵ Based on this one study, the WWC considers the extent of evidence for *Reading Edge* on adolescent readers to be small for one domain: comprehension. Three other domains are not reported in this intervention report. (See the Effectiveness Summary for further description of all domains.)

Effectiveness

Reading Edge was found to have no discernible effects on comprehension for adolescent readers.

Table 1. Summary of findings⁶

Outcome domain	Rating of effectiveness	Improvement index (percentile points)		Number of studies	Number of students	Extent of evidence
		Average	Range			
Comprehension	No discernible effects	+5	na	1	405	Small

na = not applicable

Program Information

Background

Reading Edge was developed by Robert Slavin and Nancy Madden in conjunction with Johns Hopkins University and is the middle school reading component of the SFA® whole-school reform model. *Reading Edge* is distributed by the Success for All Foundation, Inc., 200 W. Towsontown Boulevard, Baltimore, MD, 21204-5200. Email: sfainfo@successforall.org. Web: <http://www.successforall.org/Middle-High/Powerful-Instruction/The-Reading-Edge-Middle-School/>. Telephone: (800) 548-4998 ext. 2372.

Program details

Reading Edge is designed to combine explicit instruction in metacognitive reading strategies with peer collaboration, regular assessment and feedback, and classroom management techniques that are intended to engage students in learning. Although the instruction varies according to the nature of the text and the challenges it presents, every *Reading Edge* lesson follows a regular pattern: Setting the Stage, Active Instruction, Teamwork, and Time for Reflection.

The program is segmented into eight levels. Instruction is delivered in daily 60-minute lessons in which students are grouped into separate classes according to their reading levels, regardless of age or grade level. Regrouping allows teachers to teach the whole class without having to break the class into multiple smaller reading groups. Students are reassessed quarterly to determine their current level of performance.

Levels 1 through 3 of the program are remedial. Level 1 focuses on fundamental reading skills with an emphasis on individual words and improving comprehension. Using a sequence of illustrated stories, level 1 lessons present phonetically regular text that becomes more difficult as students master new skills. Levels 2 and 3 focus on fluency, comprehension skills and strategies, vocabulary, word recognition, and writing. These levels use simple fiction, nonfiction, and stories translated to scripts to teach basic decoding skills and improving reading fluency. Levels 4 through 8 add story structure, literary techniques, purposes for writing, and the writing process. These levels use short stories, novels, poetry, and nonfiction.

Professional development for *Reading Edge* teachers is provided in a two-day initial training that focuses on instructional strategies, routines, assessments, and cooperative-learning strategies, and provides support for administration and district personnel. Ongoing coaching and support are provided in quarterly coaching visits, follow-up telephone meetings, and informal telephone support for all staff members.

Cost

The *Reading Edge* program is built around classroom sets of nonfiction and fiction trade books. Schools purchase these separately from commercial vendors. Teacher manuals provide 40 weeks worth of detailed lesson plans, tests, homework, and vocabulary lists for each reading level. For details on specific product pricing, contact the program developer, the Success for All Foundation.

Research Summary

Thirty-seven studies reviewed by the WWC investigated the effects of *Reading Edge* on adolescent readers. One study (Chamberlain, Daniels, Madden, & Slavin, 2007) is a randomized controlled trial that meets WWC evidence standards without reservations. That one study is summarized in this report. The remaining 36 studies do not meet either WWC eligibility screens or evidence standards. (See references beginning on p. 5 for citations for all 37 studies.)

Summary of study meeting WWC evidence standards without reservations

Chamberlain et al. (2007) conducted a randomized controlled trial that examined the effects of *Reading Edge* on sixth-grade students attending two rural middle schools, one in Florida and one in West Virginia. The study used a two-stage random assignment design. All sixth-grade teachers were randomly assigned to teach either the *Reading Edge* program or the normal reading curriculum. All incoming sixth-grade students at the two schools were stratified by state reading assessment levels, gender, and ethnicity, and then randomly assigned to intervention or comparison teachers. The WWC based its effectiveness ratings on findings from comparisons of 203 students in the *Reading Edge* group and 202 students in the comparison group. The study reported student outcomes after approximately seven to nine months of program implementation.

Summary of studies meeting WWC evidence standards with reservations

No studies of *Reading Edge* meet WWC evidence standards with reservations.

Table 2. Scope of reviewed research

Grade	6
Delivery method	Small group/Whole class
Program type	Curriculum
Studies reviewed	37
Meets WWC standards without reservations	1 study
Meets WWC standards with reservations	0 studies

Effectiveness Summary

The WWC review of interventions for Adolescent Literacy addresses student outcomes in four domains: alphabets, reading fluency, comprehension, and general literacy achievement. The one study that contributes to the effectiveness rating in this report covers one domain: comprehension. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *Reading Edge* on adolescent readers.⁷ For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 15.

Summary of effectiveness for the comprehension domain

Chamberlain et al. (2007) did not find a statistically significant positive effect of *Reading Edge* on the total Gates-MacGintie Reading Test score for students in grade 6. The effect was not large enough to be considered substantively important according to WWC criteria (i.e., an effect size of at least 0.25).

Thus, for the comprehension domain, one study showed indeterminate effects. This results in a rating of no discernible effects, with a small extent of evidence.

Table 3. Rating of effectiveness and extent of evidence for the comprehension domain

Rating of effectiveness	Criteria met
No discernible effects <i>No affirmative evidence of effects.</i>	The review of <i>Reading Edge</i> in the comprehension domain had one study showing indeterminate effects.
Extent of evidence	Criteria met
Small	The review of <i>Reading Edge</i> in the comprehension domain was based on one study that included two schools and 405 students.

References

Study that meets WWC evidence standards without reservations

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 Research Journal*, 2(1), 1–21.

Additional sources:

Chamberlain, A., Daniels, C., Madden, N. A., & Slavin, R. E. (2009). A randomized evaluation of the Success for All middle school reading program. In D. L. Hough (Ed.), *Middle grades research: Exemplary studies linking theory to practice* (pp. 21–41). Charlotte, NC: Information Age Publishing.

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

Studies that do not meet WWC evidence standards

Borman, G. D., & Hewes, G. M. *The long-term effects and cost-effectiveness of Success for All*. Baltimore, MD: Success for All Foundation. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

Evans-Andris, M., & Usui, W. M. (2008). Comprehensive school reform & student achievement in Kentucky middle schools. *Middle Grades Research Journal*, 3(2), 51–69. The study does not meet WWC evidence standards because the measures of effectiveness cannot be attributed solely to the intervention—the intervention was combined with another intervention.

Gines, B. E. (2009). Comparing two comprehensive reform models: Their effect on student reading achievement. *Dissertation Abstracts International*, 68(8A). The study does not meet WWC evidence standards because the measures of effectiveness cannot be attributed solely to the intervention—there was only one unit assigned to one or both conditions.

Ross, S. M., Nunnery, J. A., Goldfeder, E., McDonald, A., Rachor, R., Hornbeck, M., & Fleischman, S. (2004). Using school reform models to improve reading achievement: A longitudinal study of Direct Instruction and Success for All in an urban district. *Journal of Education for Students Placed at Risk*, 9(4), 357–388. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

Taylor, G. B. (2002). *The impact of Literacy Collaborative and Success for All on the reading achievement of third and fifth grade students*. Unpublished doctoral dissertation, University of Georgia, Athens. The study does not meet WWC evidence standards because the measures of effectiveness cannot be attributed solely to the intervention—there was only one unit assigned to one or both conditions.

Studies that are ineligible for review using the Adolescent Literacy Evidence Review Protocol

American Federation of Teachers. (1999). *Building on the best, learning from what works: Five promising remedial reading intervention programs*. Washington, DC: Author. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Additional source:

American Federation of Teachers. (1998). *Building on the best, learning from what works: Seven promising reading and English language arts programs*. Washington, DC: Author.

Briggs, K. L., & Clark, C. (1997). *Reading programs for students in the lower elementary grades: What does the research say?* Austin: Texas Center for Educational Research. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Bullas, A. C. R. (2009). Family perceptions of parent involvement programs and their influence on student achievement. *Dissertation Abstracts International*, 70(4A). The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.

- Chambers, B., Cheung, A., Madden, N. A., Slavin, R. E., & Gifford, R. (2009). Embedded multimedia: Using video to enhance reading outcomes in Success for All. In A. G. Bus & S. B. Neuman (Eds.), *Multimedia and literacy development: Improving achievement for young learners* (pp. 213–223). New York: Routledge. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Chambers, B., Slavin, R. E., Madden, N. A., Ambrami, P. C., Tucker, B. J., & Cheung, A. (2009). Computer-assisted tutoring in Success for All: Two studies of reading outcomes for first graders. In A. G. Bus & S. B. Neuman (Eds.), *Multimedia and literacy development: Improving achievement for young learners* (pp. 224–237). New York: Routledge. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Chambers, B., Slavin, R. E., Madden, N. A., Cheung, A., & Gifford, R. (2005). *Enhancing Success for All for Hispanic students: Effects on beginning reading achievement*. Baltimore, MD: Success for All Foundation. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- 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. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Crowe, E. C., Connor, C. M., & Petscher, Y. (2009). Examining the core: Relations among reading curricula, poverty, and first through third grade reading achievement. *Journal of School Psychology*, 47(3), 187–214. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Dion, E., Brodeur, M., Gosselin, C., Campeau, M., & Fuchs, D. (2010). Implementing research-based instruction to prevent reading problems among low-income students: Is earlier better? *Learning Disabilities Research & Practice*, 25(2), 87–96. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Gross, B., Booker, T. K., & Goldhaber, D. (2009). Boosting student achievement: The effect of comprehensive school reform on student achievement. *Educational Evaluation and Policy Analysis*, 31(2), 111–126. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Gyovai, L. K., Cartledge, G., Kourea, L., Yurick, A., & Gibson, L. (2009). Early reading intervention: Responding to the learning needs of young at-risk English language learners. *Learning Disability Quarterly*, 32(3), 143–162. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Hurley, E. A., Chamberlain, A., Slavin, R. E., & Madden, N. A. (2001). Effects of Success for All on TAAS reading scores. *Phi Delta Kappan*, 82(10), 750–756. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Jackson, W. D. (2008). An investigation of the impact of the Success for All whole-school reform model on the Elementary School Proficiency Assessment and the New Jersey Assessment of Skills and Knowledge in an urban district. *Dissertation Abstracts International*, 69(2), 490. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample includes less than 50% general education students.
- Lockwood, A. T., & Secada, W. G. (1999). *Transforming education for Hispanic youth: Exemplary practices, programs, and schools*. Washington, DC: National Clearinghouse for Bilingual Education. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Malmgren, K. W., & Trezek, B. J. (2009). Literacy instruction for secondary students with disabilities. *Focus on Exceptional Children*, 41(6), 1–12. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Manset, G., St. John, E. P., Simmons, A., Michael, R., Bardzell, J., Hodges, D., . . . Gordon, D. (1999). *Indiana's early literacy intervention grant program impact study for 1997–98*. Bloomington: Indiana Education Policy Center, Smith Center for Research in Education. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Additional source:

Manset, G., St. John, E. P., & Simmons, A. B. (2000). *Progress in early literacy: Summary evaluation of Indiana's early literacy intervention grant program, 1997–2000*. Bloomington: Indiana Education Policy Center, Smith Center for Research in Education.

Marsh, J., Hamilton, L., & Gill, B. (2008). Assistance and accountability in externally managed schools: The case of Edison Schools, Inc. *Peabody Journal of Education*, 83(3), 423–458. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Reschly, A. L. (2010). Reading and school completion: Critical connections and Matthew effects. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 26(1), 67–90. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Rissman, L. (2005). *The Reading Edge*. Tallahassee: Florida Center for Reading Research. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Sadler, C., & Sugai, G. (2009). Effective behavior and instructional support: A district model for early identification and prevention of reading and behavior problems. *Journal of Positive Behavior Interventions*, 11(1), 35–46. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Simon, J. (2011). *A cost-effectiveness analysis of early literacy interventions*. Unpublished doctoral dissertation, Columbia University, New York. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Slavin, R. E., Chamberlain, A., & Daniels, C. (2007). Preventing reading failure. *Educational Leadership*, 65(2), 22–27. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly*, 43(3), 290–322. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2009). *Effective programs for struggling readers: A best evidence synthesis*. Baltimore, MD: Johns Hopkins University, Center for Data-Driven Reform in Education. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Slavin, R. E., & Madden, N. A. (2004). *Success for All/Roots & Wings: Summary of research on achievement outcomes* (Report No. 41). Baltimore, MD: Center for Research on the Education of Students Placed At Risk (CRESPAR). The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Slavin, R. E., & Madden, N. A. (2006). *Success for All: Summary of research on achievement outcomes*. Baltimore, MD: Center for Data-Driven Reform in Education at Johns Hopkins University. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Slavin, R. E., & Madden, N. A. (2010). Success for All: Prevention and early intervention in school wide reform. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 434–445). New York: Routledge. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

- Success for All Foundation. (2006). *Summary of research on the Success for All reading programs*. Baltimore, MD: Author. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Success for All Foundation. (2008). *Randomized research proves Success for All raises reading achievement*. Baltimore, MD: Author. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Tivnan, T., & Hemphill, L. (2005). Comparing four literacy reform models in high-poverty schools: Patterns of first-grade achievement. *Elementary School Journal*, 105(5), 419–441. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Uhry, J. K., & Clark, D. B. (2004). *Dyslexia: Theory & practice of instruction* (3rd ed.). Austin, TX: Pro-Ed. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Appendix A: Research details for Chamberlain et al. (2007).

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 Research Journal*, 2(1), 1–21.

Table A. Summary of findings

Meets WWC evidence standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index (percentile points)	Statistically significant
Comprehension	405 students	+5	No

Setting The study was conducted in two rural middle schools, one in Florida and the other in West Virginia. Both schools served high-poverty, majority-White populations. In the Florida school, 69% of the students qualified for free lunch, and 75% of the students were White. In the West Virginia school, 50% of the students qualified for free lunch, and 95% of the students were White.

Study sample Recruitment of schools occurred in 2005, and the study was conducted during the 2005–06 school year. The study used a two-stage random assignment design. All sixth-grade teachers in the study schools were randomly assigned to teach either the *Reading Edge* program (intervention) or the normal reading curriculum (comparison). All incoming sixth-grade students at the two schools were stratified by state reading assessment levels, gender, and ethnicity, and then randomly assigned to intervention or comparison teachers. The analysis sample used in the study included 203 students in the *Reading Edge* group and 202 students in the comparison group. The study used a delayed treatment design, whereby comparison teachers received the *Reading Edge* materials and training in the fall of 2008 (after the study’s data collection period).

Intervention group Students in the intervention group received daily reading instruction from teachers who were given materials and training to implement the *Reading Edge* program. Students were assessed and then grouped in classes by their reading level. They were reassessed quarterly to allow for movement to other group levels and for targeted instruction. The intervention as implemented in this study differs slightly from the implementation of *Reading Edge* in the context of the SFA® whole-school reform. Although most SFA® *Reading Edge* schools use a full-time, school-based facilitator, these research sites did not have facilitators. In addition, most schools involve a large percentage of their teaching staff in the program; these research sites had a limited number of participating teachers. Finally, most schools have one to three common reading periods for regrouping per day; these research sites did not have common reading periods. The study reported student outcomes after approximately seven to nine months of program implementation.

Comparison group Students in the comparison group received daily reading instruction using “business-as-usual” methods and materials. In both schools, comparison students were sorted into comparison classes based on fifth-grade reading scores. The comparison group teachers in the Florida school used *Corrective Reading* for students with low reading skills and *Scholastic Read XL* for other students. The comparison group teachers in the West Virginia school used *Scholastic Read 180* and the *Scholastic Read Aloud Anthology* for students with low reading skills, Larry Bell’s “Twelve Powerful Words” in all reading classes, and a teacher-designed “Literacy Circle” for small groups of more proficient students.

Outcomes and measurement

For the pretest, students took the Scholastic Reading Inventory in the fall of 2005. For the posttest, the students took the Gates-MacGinitie Reading Test (GMRT) at the end of the school year, in the spring of 2006. The study reports total GMRT scores, which are the average of scores on two subscales: Comprehension and Vocabulary. For a more detailed description of the posttest outcome measures, see Appendix B. Findings for the total GMRT can be found in Appendix C. Subtest findings can be found in Appendix D. The study also measured classroom instructional elements using a classroom observation tool based on the Classroom Assessment Scoring System (CLASS), but these measures did not meet inclusion criteria as outcomes for the Adolescent Literacy Topic Area. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Training was provided to teachers for the implementation of *Reading Edge*. No details are given about the extent and timing of the training.

Appendix B: Outcome measures for each domain

Comprehension	
<i>Gates-MacGinitie Reading Test (GMRT), total score</i>	The GMRT is used to assess a student's decoding, vocabulary, and passage comprehension skills. The total score is an average of two components that independently assess reading vocabulary and comprehension skills (as cited in Chamberlain et al., 2007).
Reading comprehension construct	
<i>GMRT, Comprehension subtest</i>	The Comprehension subtest measures each student's ability to read and understand different types of prose. The test contains 11 passages of various lengths and subjects and 48 questions (as cited in Chamberlain et al., 2007).
Vocabulary development construct	
<i>GMRT, Vocabulary subtest</i>	The Vocabulary subtest measures each student's reading vocabulary by asking the student to choose one word or phrase that means most nearly the same as a presented word. The test contains 45 questions (as cited in Chamberlain et al., 2007).

Appendix C: Findings included in the rating for the comprehension domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Chamberlain et al. (2007)^a								
<i>GMRT total score</i>	Grade 6	405 students	57.93 (16.15)	55.76 (16.10)	2.17	0.13	+5	0.06
Domain average for comprehension (Chamberlain et al., 2007)						0.13	+5	Not statistically significantly
Domain average for comprehension across all studies						0.13	+5	Not statistically significantly

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention. The statistical significance of each study’s domain average was determined by the WWC. For Chamberlain et al. (2007), the effect was not statistically significant ($p > 0.05$). GMRT= Gates-MacGinitie Reading Test.

^a For Chamberlain et al. (2007), no corrections for clustering or multiple comparisons and no difference-in-differences adjustment were needed. The p-value presented here was reported in the original study. The means shown in this table are the adjusted posttest means reported in the study.

Appendix D: Supplemental subtest findings for the comprehension domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Chamberlain et al. (2007)^a								
<i>GMRT–Comprehension</i>	Grade 6	405 students	29.87 (9.14)	28.90 (8.92)	0.97	0.11	+4	0.16
<i>GMRT–Vocabulary</i>	Grade 6	405 students	28.05 (7.92)	26.85 (8.35)	1.20	0.15	+6	0.05

Table Notes: The supplemental findings presented in this table are additional subtest findings from the study in this report that do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention. GMRT = Gates-MacGinitie Reading Test.

^a For Chamberlain et al. (2007), a correction for multiple comparisons was needed and resulted in significance levels that differ from those in the original study. The p-values presented here were reported in the original study. Due to the multiple comparisons adjustment, the p-value of 0.05 for the Vocabulary subtest of the GMRT was higher than the critical p-value of 0.025 for statistical significance; therefore, the WWC does not find the result to be statistically significant. The means shown in this table are the adjusted posttest means reported in the study.

Endnotes

¹ The descriptive information for this program was obtained from a publicly available source: the program's website (<http://www.successforall.org/Middle-High/Powerful-Instruction/The-Reading-Edge-Middle-School/>, downloaded June 2011). The WWC requests developers review the program description sections for accuracy from their perspective. The program description was provided to the developer in July 2011; however, the WWC received no response. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review. The literature search reflects documents publicly available by August 2011.

² *Success for All (SFA)*[®] is a whole-school reform model that includes reading, writing, and oral language development programs for students in prekindergarten through eighth grade.

³ Studies of *Reading Edge* conducted in grade 3 and lower are outside of the scope of the Adolescent Literacy review protocol.

⁴ The studies in this report were reviewed using WWC Evidence Standards, version 2.1, as described in the Adolescent Literacy review protocol, version 2.0.

⁵ The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

⁶ For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p.15. These improvement index numbers show the average and range of student-level improvement indices for all findings across the study.

⁷ The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Chamberlain et al. (2007), a correction for multiple comparisons was needed for the subscale scores presented in Appendix C, so the significance levels may differ from those reported in the original study.

⁸ The WWC computes an average effect size as a simple average of the effect sizes across all individual findings within the study domain. The average effect size across the two subtest measures corresponds to the GMRT total score effect size.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, June). *Adolescent Literacy intervention report: Reading Edge*. Retrieved from <http://whatworks.ed.gov>.

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC evidence standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC evidence standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
Positive effects	Two or more studies show statistically significant positive effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important negative effects.
Potentially positive effects	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
Mixed effects	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows a statistically significant or substantively important negative effect, but no more such studies than the number showing a statistically significant or substantively important positive effect, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
Potentially negative effects	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
Negative effects	Two or more studies show statistically significant negative effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important positive effects.
No discernible effects	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

Glossary of Terms

Attrition	Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.
Clustering adjustment	If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.
Confounding factor	A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.
Design	The design of a study is the method by which intervention and comparison groups were assigned.
Domain	A domain is a group of closely related outcomes.
Effect size	The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.
Eligibility	A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.
Equivalence	A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.
Extent of evidence	An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 15.
Improvement index	Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from -50 to +50.
Multiple comparison adjustment	When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.
Quasi-experimental design (QED)	A quasi-experimental design (QED) is a research design in which subjects are assigned to intervention and comparison groups through a process that is not random.
Randomized controlled trial (RCT)	A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into intervention and comparison groups.
Rating of effectiveness	The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 15.
Single-case design	A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.
Standard deviation	The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.
Statistical significance	Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).
Substantively important	A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the [WWC Procedures and Standards Handbook \(version 2.1\)](#) for additional details.