What Works Clearinghouse

Student Team Reading and Writing

Report Contents

Program Description[1]

Student team reading and writing[2] refers to two cooperative learning programs for secondary students included in this intervention report: (1) *Student Team Reading and Writing* and (2) *Student Team Reading*. The *Student Team Reading and Writing* program (Stevens, 2003) is an integrated approach to reading and language arts for early adolescents. The program incorporates (1) cooperative learning classroom processes; (2) a literature anthology for high-interest reading material; (3) explicit instruction in reading comprehension; (4) integrated reading, writing, and language arts instruction; and (5) a writing process approach to language arts. *Student Team Reading* (Stevens, 1989; Stevens & Durkin, 1992) comprises the reading part of *Student Team Reading and Writing* and consists of two principal elements: (1) literature-related activities (including partner reading, treasure hunts, word mastery, story retelling, story-related writing, and quizzes) and (2) direct instruction in reading comprehension strategies (such as identifying main ideas and themes, drawing conclusions, making predictions, and understanding figurative language). The writing part of the *Student Team Reading and Writing* program includes selection-related writing. As part of the two programs that are the focus of this report, students work in heterogeneous learning teams, and activities are designed to follow a regular cycle that involves teacher presentation, team practice, independent practice, peer pre-assessment, and individual assessments that form the basis for team scores. The cooperative learning teams used in the programs are intended to engage students in academic interactions.

Research[3]

Two studies of *student team reading and writing* that fall within the scope of the Adolescent Literacy review protocol meet What Works Clearinghouse (WWC) evidence standards with reservations. The two studies included more than 5,200 adolescent learners from grades 6 through 8 in urban middle schools in the eastern United States. Based on these two studies, the WWC considers the extent of evidence for *student team reading and writing* on adolescent learners to be medium to large for the comprehension domain and small for the general literacy achievement domain. The two studies that meet WWC evidence standards with reservations did not examine the effectiveness of student team reading and writing on adolescent learners in the alphabetics and reading fluency domains.

Effectiveness

Student team reading and writing was found to have potentially positive effects on comprehension and no discernible effects on general literacy achievement for adolescent learners.

Table 1. Summary of findings[4]

Outcome domain	Rating of effectiveness	Improvement index *(percentile points)*		Number of studies	Number of students	Extent of evidence
		Average	Range			
Alphabetics	na	na	na	na	na	na
Reading fluency	na	na	na	na	na	na
Comprehension	Potentially positive effects	+7	–1 to +13	2	5,209	Medium to large
General literacy achievement	No discernible effects	+8	0 to +16	1	3,986	Small

na = not applicable

Program Information

Background

Robert Stevens developed *Student Team Reading* and *Student Team Reading and Writing* in 1989 and 1992, respectively.

Student Team Reading and Writing and *Student Team Reading* developer: Robert Stevens. Address: The Pennsylvania State University, 202 CEDAR Building, University Park, PA 16802. Email: rjs15@psu.edu. Web: http://www.ed.psu.edu/educ/espse/edpsych/people/robert-stevens. Telephone: 814-863-2417.

Program details

According to the developers, the program was originally developed and used in urban classrooms and since then has been implemented with culturally and linguistically diverse populations. This review includes one study each of *Student Team Reading and Writing* and *Student Team Reading* programs.

The *Student Team Reading and Writing and Student Team Reading* programs use a literature anthology as the basis for the reading material. Teachers introduce narratives from the literature anthology with discussions of relevant background knowledge, genre, and vocabulary. Students work in cooperative learning teams of four to five as they read and receive rewards for working well both as an individual and as a group member. Through the use of literature, students learn about different genres of writing and become more familiar with famous, well-published authors (e.g., O. Henry, Langston Hughes) whom they could choose to read more extensively on their own.

Typical activities include the following:

1. *Partner Reading*. Students first read silently, then take turns reading orally with a partner.
2. *Treasure Hunts*. Questions guide student reading, requiring them to search and think to generate text-supported answers.
3. *Word Mastery*. Students practice saying new vocabulary words with their partners. They then use those words in story-related writing. (Story-related writing is part of the *Student Team Reading and Writing* program).
4. *Story Retelling*. Students summarize stories in their own words.
5. *Story-Related Writing*. Students write in responses to prompts about their reading. This activity is part of the *Student Team Reading and Writing* program.
6. *Extension Activities*. Students complete cross-curricular research, fine arts, dramatics, and media activities as they explore themes in the stories/books.
7. *Tests*. Students take tests on comprehension, word meaning, and word pronunciation.

8. *Explicit Instruction of Comprehension Strategies.* Teachers model and guide students in the use of comprehension strategies.

Cost

Cost information associated with the *Student Team Reading and Writing* and *Student Team Reading* programs is available from the developer.

Research Summary

Four studies reviewed by the WWC Adolescent Literacy Topic Area investigated the effects of *student team reading and writing* on adolescent learners. Two studies (Stevens, 2003; Stevens & Durkin, 1992) are quasi-experimental designs that meet WWC evidence standards with reservations. The remaining two studies do not meet either WWC eligibility screens or evidence standards. (See references beginning on page 7 for citations for all four studies.)

Table 2. Scope of reviewed research

Grade	6, 7, 8
Delivery method	Small group
Program type	Practice
Studies reviewed	4
Meets WWC standards	0 studies
Meets WWC standards with reservations	2 studies

Summary of studies meeting WWC evidence standards without reservations

No studies of *Student Team Reading and Writing* meet WWC evidence standards without reservations.

Summary of studies meeting WWC evidence standards with reservations

Stevens (2003) conducted a quasi-experiment that examined the effects of the *Student Team Reading and Writing* curriculum on students in grades 6 through 8 attending five middle schools in a large urban school district in the eastern United States. The investigator matched two treatment schools with three comparison schools on students' ethnicity, socioeconomic background, and academic achievement in reading and language arts. The WWC based its effectiveness ratings on findings from comparisons of 1,798 students in two schools who received *Student Team Reading and Writing* and 2,188 comparison group students in three schools who received regular instruction. The study reported students' outcomes after nine months of program implementation.

Stevens and Durkin (1992) conducted a quasi-experiment that examined the effects of the *Student Team Reading* curriculum on students in grade 6 attending six middle schools in an urban school district in Maryland. Twenty experimental classes were matched on California Achievement Test reading scores with 34 comparison classes. The WWC based its effectiveness ratings on findings from comparisons of 455 students in three schools who received *Student Team Reading* and 768 comparison group students in three schools who received regular instruction. The study reported students' outcomes after nine months of program implementation.

Effectiveness Summary

The WWC review of interventions for Adolescent Literacy addresses student outcomes in four domains: alphabetics, reading fluency, comprehension, and general literacy achievement. The two studies that influence the findings in this report cover two domains: alphabetics and general literacy achievement. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *student team reading and writing* on adolescent learners. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see Appendix D.

Summary of effectiveness for the comprehension domain

Two studies reported findings in the comprehension domain.

Stevens (2003) reported statistically significant positive effects of *Student Team Reading and Writing* on the Reading Comprehension and Vocabulary subtests of the California Achievement Test (CAT) for students in grades 6 through 8. According to WWC calculations, the effects were not statistically significant (when adjusted for clustering), but the average effect across the two measures was large enough to be considered substantively important according to WWC criteria (i.e., an effect size of at least 0.25).[5]

Stevens and Durkin (1992) found a statistically significant positive effect of *Student Team Reading* on the Reading Comprehension subtest of the CAT for students in grade 6, but they did not find statistically significant effects on the CAT Vocabulary subtest. The WWC-calculated effects were not statistically significant, and the average effect across the two measures was not large enough to be considered substantively important according to WWC criteria.

Thus, for the comprehension domain, one study showed substantively important positive effects and one study showed indeterminate effects, with a medium to large extent of evidence.

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

Rating of effectiveness	Criteria met
Potentially positive effects *Evidence of a positive effect with no overriding contrary evidence*	The review of *student team reading and writing* had one study showing substantively important positive effects and one study showing indeterminate effects.
Extent of evidence	**Criteria met**
Medium to large	The review of *student team reading and writing* was based on two studies which included 11 schools and 5,209 students.

Summary of effectiveness for the general literacy achievement domain

One study reported findings in the general literacy achievement domain.

Stevens (2003) found a statistically significant positive effect of *Student Team Reading and Writing* on the CAT Language Expression subtest for students in grades 6 through 8 but did not find statistically significant effects on the CAT Language Mechanics subtest. According to WWC calculations, the effect on the Language Expression subtest was not statistically significant (when adjusted for clustering), and the average effect across the two measures was not large enough to be considered substantively important according to WWC criteria.

Thus, for the general literacy achievement domain, one study showed indeterminate effects, with a small extent of evidence.

Table 4. Rating of effectiveness and extent of evidence for the general literacy achievement domain

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects*	The review of *Student Team Reading and Writing* had one study showing indeterminate effects.
Extent of evidence	**Criteria met**
Small	The review of *Student Team Reading and Writing* was based on one study that included five schools and 3,986 students.

References

Studies that meet WWC evidence standards with reservations

Stevens, R. J. (2003). Student team reading and writing: A cooperative learning approach to middle school literacy instruction. *Educational Research and Evaluation, 9*(2), 137–160.

Additional sources:

Stevens, R. J. (2006). Integrated reading and language arts instruction. *RMLE Online: Research in Middle Level Education, 30*(3), 1–12.

Stevens, R. J., & Durkin, S. (1992). *Using student team reading and student team writing in middle schools: Two evaluations. Part II. Student team reading and student team writing: An evaluation of a middle school reading and writing program* (pp. 1–11). Baltimore, MD: Center for Research on Effective Schooling for Disadvantaged Students.

Stevens, R. J., & Durkin, S. (1992). *Using student team reading and student team writing in middle schools: Two evaluations. Part I. Student team reading: A cooperative learning approach to middle school reading instruction* (pp. 1–10). Baltimore, MD: Center for Research on Effective Schooling for Disadvantaged Students.

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

Mac Iver, D. J., & Plank, S. B. (1996). *The talent development middle school. Creating a motivational climate conducive to talent development in middle schools: Implementation and effects of Student Team Reading* (Report No. 4). Baltimore, MD: Howard University and Johns Hopkins University Center for Research on the Education of Students Placed At Risk. The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.

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 not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Appendix A.1: Research details for Stevens (2003)

Stevens, R. J. (2003). Student team reading and writing: A cooperative learning approach to middle school literacy instruction. *Educational Research and Evaluation, 9*(2), 137–160.

Table A1. Summary of findings

Meets WWC evidence standards with reservations

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Comprehension	5 schools/3,986 students	+11	No
General literacy achievement	5 schools/3,986 students	+8	No

Setting

The study took place in five middle schools in the Baltimore City Public School System. The study school population was predominantly minority students (approximately 80%), and approximately 67% of students in the study schools received free or reduced-price lunch.

Study sample

This study is a quasi-experiment conducted in five urban middle schools. Two treatment schools volunteered to implement the intervention, and three schools, matched on academic achievement in reading and language arts on the California Achievement Test, served as comparison schools. The author also matched the schools on ethnicity and socioeconomic background. Participants were sixth-, seventh-, and eighth-grade students. The study's analytic sample included 1,798 students in 72 treatment classrooms and 2,188 students in 88 comparison classrooms.

Intervention group

The study reported students' outcomes after nine months of *Student Team Reading and Writing* implementation. The intervention components included (1) cooperative learning classroom processes; (2) a literature anthology for high-interest reading material; (3) explicit instruction in reading comprehension; (4) integrated reading, writing, and language arts instruction; and (5) a writing process approach to language arts. The reading part of the program consisted of three elements: (1) literature-related activities, (2) direct instruction in reading comprehension strategies, and (3) selection-related writing. In all of these activities, students worked in heterogeneous learning teams. All activities followed a regular cycle that involved teacher presentation, team practice, independent practice, peer pre-assessment, and individual accountability. Cooperative learning teams were used in instructional activities on a daily basis.

Comparison group

The teachers in the comparison schools used traditional instructional methods. Students went to different teachers for reading and English. The reading teachers used a basal reading series and related materials (e.g., workbooks). The English teachers used an English literature anthology for their literature component and an English grammar textbook for the language arts component. The comparison teachers did not use cooperative learning processes in their instructional activities.

Outcomes and measurement

For both the pretest and posttest, students took the California Achievement Test (CAT). Four subtests were used in the study: (1) Reading Vocabulary, (2) Reading Comprehension, (3) Language Mechanics, and (4) Language Expression. The pretests were given the spring before the study began; the posttests were given the following May near the end of the study. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

The teachers in the treatment schools were trained in *Student Team Reading and Writing* during their summer vacation for five half-day (three-hour) sessions during one week. The training consisted of an explanation of the processes and the rationale behind them. During the training, teachers participated in a simulation of major components of the program. The teachers also were given a detailed manual that described each of the components. During the first three months of implementation, the researchers observed and gave feedback to the teachers as they implemented the program. They also met with teachers during and after school, often attending meetings of the reading and language arts department. At these meetings, teachers' questions and problems were discussed in order to resolve any problems they were having and to use their feedback to improve the program.

Appendix A.2: Research details for Stevens & Durkin (1992)

Stevens, R. J., & Durkin, S. (1992). *Using student team reading and student team writing in middle schools: Two evaluations. Part I. Student team reading: A cooperative learning approach to middle school reading instruction* (pp. 1–10). Baltimore, MD: Center for Research on Effective Schooling for Disadvantaged Students.

Table A2. Summary of findings

Meets WWC evidence standards with reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Comprehension	6 schools/1,223 students	+2	No

Setting

The study took place in six middle schools in an urban school district in Maryland. The student populations in the schools ranged from 27% to 99% minority (median of 74.5%) and from 38% to 77% received free or reduced-price lunch (median of 57.5%).

Study sample

This study is a quasi-experiment conducted in six urban middle schools during the 1989–90 academic year. Classes in the three treatment schools were matched with classes in the three comparison schools on California Achievement Test (CAT) total reading pretest scores. Participants were sixth-grade students. The study's analytic sample included 455 students in 20 treatment classrooms and 768 students in 34 comparison classrooms.

Intervention group

In the intervention schools, *Student Team Reading* was implemented for one full academic year and included two major components: (1) literature-based activities (including partner reading, treasure hunts, word mastery, story retelling, story-related writing, and quizzes) and (2) explicit instruction in comprehension strategies (such as identifying main ideas and themes, drawing conclusions, making predictions, and understanding figurative language). The program used a combination of teacher-directed instruction and cooperative learning in heterogeneous teams. Teams were given rewards and recognition based on performance and improvement of each team member.

Comparison group

Comparison group teachers used traditional methods and curriculum materials. In reading, they often used basal series and focused instruction on isolated skills. Students read silently and aloud (with one student reading while the rest of the class follows along). Most seatwork time was spent on independent class work completing worksheet activities and practicing reading skills; at times, students read silently; at other times, they read orally in turns. Students typically did little to no extended writing that was related to reading activities.

Outcomes and measurement

For both the pretest and posttest, students took the CAT. Two CAT subtests were used in the study: Reading Vocabulary and Reading Comprehension. The fifth-grade scores were used as the pretest data; the sixth-grade scores were used as the posttest data. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Intervention group teachers took part in three half-day training sessions that included how to implement the classroom processes and the rationale behind the processes. During the training session, trainers acted as the "teachers," and the teachers acted as the students. Teachers also received a detailed manual of the *Student Team Reading* program, curriculum materials, and textbooks. During the school year, teachers also received coaching and took part in periodic after-school meetings to provide feedback/discuss implementation questions. Teachers were monitored by *Student Team Reading* staff four times a week over a six-week period.

Appendix B: Outcome measures for each domain

Comprehension	
Vocabulary development construct	
California Achievement Test (CAT) Vocabulary subtest	The California Achievement Test is a norm- and criterion-referenced annual test. The Vocabulary subtest contains 20 items measuring same-meaning, opposite-meaning words; multi-meaning words; words in context (selecting the appropriate word to complete the sentence); and the meaning of affixes (as cited in Stevens, 2003; Stevens & Durkin, 1992).
Reading comprehension construct	
California Achievement Test (CAT) Reading Comprehension subtest	The California Achievement Test is a norm- and criterion-referenced annual test. The Reading Comprehension subtest is administered to grades 1 through 12 and focuses on students' use of reading comprehension strategies. Students read a multiple-paragraph passage and then answer (factual and summarization) questions about the passage. Passages reflect a wide range of narrative, expository, contemporary, and traditional texts. The test measures information recall, meaning construction, form analysis, and meaning evaluation of different selections (as cited in Stevens, 2003; Stevens & Durkin, 1992).
General literacy achievement	
California Achievement Test (CAT) Language Mechanics subtest	The Language Mechanics and Language Expression subtests of the California Achievement Test work together to measure a broad range of language and writing skills, including the ability to apply standard usage and writing conventions and to develop effective sentences and paragraphs. The Language Mechanics subtest contains 20 items that measure skills in the mechanics of capitalization and punctuation. Editing skills are measured in the context of passages presented in various formats (for grades 4–12). Students read sentences and passages and identify errors (or report "no errors") that occur in the passage (as cited in Stevens, 2003).
California Achievement Test (CAT) Language Expression subtest	The Language Mechanics and Language Expression subtests of the California Achievement Test work together to measure a broad range of language and writing skills, including the ability to apply standard usage and writing conventions and to develop effective sentences and paragraphs. The Language Expression subtest contains 20 items that measure skills in language usage and sentence structure. The items measure skills in the use of various parts of speech, formation and organization of sentences and paragraphs, and writing for clarity (for grades 4–12). Language Expression questions ask students to read a passage and identify appropriate phrases or words to fit particular parts of the passage (as cited in Stevens, 2003).

Appendix C.1: 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	
Stevens, 2003[a]								
CAT Reading Vocabulary	Grades 6–8	5/3,986	0.17 (0.71)	–0.16 (0.72)	0.33	0.33	+13	< 0.05
CAT Reading Comprehension	Grades 6–8	5/3,986	0.12 (0.66)	–0.13 (0.73)	0.25	0.25	+10	< 0.05
Average for comprehension (Stevens, 2003)						**0.29**	**+11**	**Not sig**
Stevens & Durkin, 1992[b]								
CAT Reading Vocabulary	Grade 6	6/1,223	18.20 (5.40)	18.30 (5.60)	–0.10	–0.02	–1	> 0.05
CAT Reading Comprehension	Grade 6	6/1,223	24.30 (7.30)	23.20 (8.10)	1.10	0.14	+6	< 0.05
Average for comprehension (Stevens & Durkin, 1992)						**0.06**	**+2**	**Not sig**
Domain average for comprehension across all studies						**0.18**	**+7**	**na**

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor 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 that 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 WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of a study's domain average was determined by the WWC. Not sig = not statistically significant. na = not applicable. CAT = California Achievement Test.

[a] For Stevens (2003), corrections for clustering and multiple comparisons were 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. The intervention and comparison group values are adjusted z-scores when the pretest was used as an adjustment.

[b] For Stevens and Durkin (1992), corrections for clustering and multiple comparisons were 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. The WWC calculated the program group mean using a difference-in-differences approach (see WWC Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and control groups) to the unadjusted control group posttest means.

Appendix C.2: Findings included in the rating for the general literacy achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Stevens, 2003[a]								
CAT Language Mechanics	Grades 6–8	5/3,986	0.00 (0.73)	0.00 (0.75)	0.00	0.00	0	> 0.05
CAT Language Expression	Grades 6–8	5/3,986	0.19 (0.72)	–0.19 (0.73)	0.38	0.38	+15	< 0.05
Domain average for general literacy achievement (Stevens, 2003)						**0.19**	**+8**	**Not sig**

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor 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 that 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 WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of a study's domain average was determined by the WWC. Not sig = not statistically significant. CAT= California Achievement Test.

[a] For Stevens (2003), corrections for clustering and multiple comparisons were 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. The intervention and comparison group values are adjusted z-scores when the pretest was used as an adjustment.

Appendix D.1: Criteria used to determine the rating of a study

Study rating	Criteria
Meets evidence standards	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets 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.

Appendix D.2: 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.

Appendix D.3: 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.

Endnotes

[1] The descriptive information for this program was obtained from publicly available sources: Stevens, 1989, 2003; Stevens and Durkin, 1992. The WWC requests developers to review the program description sections for accuracy from their perspective. The program description was provided to the developer in February 2011. 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] This review focuses on *Student Team Reading and Writing* and *Student Team Reading*, but it uses the general (lowercase) term *student team reading and writing* to encompass both programs.

[3] The studies in this report were reviewed using WWC Evidence Standards, Version 2.0, as described in the Adolescent Literacy review protocol Version 2.0. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

[4] For criteria used in the determination of the rating of effectiveness and extent of evidence, see Appendix D. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies.

[5] The WWC computes an average effect size as a simple average of the effect sizes across all individual findings within the study domain.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2011, November). *Adolescent Literacy intervention report: Student team reading and writing.* Retrieved from http://whatworks.ed.gov.

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 treatment 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 standardized measure of the magnitude of an effect that is comparable across studies and outcomes.

Eligibility A study is eligible for review if it falls within the scope of the review protocol and uses a causal design (RCT or QED).

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 Appendix D.3.

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 treatment 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 treatment 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 Appendix D.2.

Standard deviation The standard deviation across all students in a group shows how dispersed the outcomes are. A measure with a small standard deviation would indicate that participants had more similar outcomes than a measure with a large standard deviation.

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 2.0) for additional details.