What Works Clearinghouse

Cooperative Integrated Reading and Composition®

Program Description[1]

Cooperative Integrated Reading and Composition® (CIRC®) is a reading and writing program for students in grades 2–6. It has three principal elements: story-related activities, direct instruction in reading comprehension, and integrated language arts/writing. Daily lessons provide students with an opportunity to practice comprehension and reading skills in pairs and small groups. Pairs of students read to each other; predict how stories will end; summarize stories; write responses to questions posed by the teacher; and practice spelling, decoding, and vocabulary. Within cooperative teams of four, students work to understand the main idea of a story and work through the writing activities linked to the story. A Spanish version of the program is available for grades 2–5.

Research[2]

One study of *Cooperative Integrated Reading and Composition®* that falls within the scope of the Beginning Reading review protocol meets What Works Clearinghouse (WWC) evidence standards without reservations, and one study meets WWC evidence standards with reservations. The two studies included approximately 700 students in grades 3 and 4 who attended elementary schools in Ohio and Pennsylvania. Based on these two studies, the WWC considers the extent of evidence for *Cooperative Integrated Reading and Composition®* on beginning readers to be medium to large for comprehension and small for general reading achievement. No studies that meet WWC evidence standards with or without reservations examined the effectiveness of *Cooperative Integrated Reading and Composition®* on beginning readers in the alphabetics and fluency domains. (See the Effectiveness Summary for further description of all domains.)

Effectiveness

Cooperative Integrated Reading and Composition® was found to have potentially positive effects on comprehension and no discernible effects on general reading achievement for beginning readers.

Table 1. Summary of findings[3]

Outcome domain	Rating of effectiveness	Improvement index *(percentile points)*		Number of studies	Number of students	Extent of evidence
		Average	Range			
Comprehension	Potentially positive effects	+12	+1 to +30	2	712	Medium to large
General reading achievement	No discernible effects	+1	na	1	320	Small

na = not applicable

Report Contents

Program Information

Background

Developed in 1983 by Robert Slavin and Nancy Madden at the Center for Social Organization of Schools at the Johns Hopkins University, *Cooperative Integrated Reading and Composition*® is distributed by the Success for All Foundation, Inc. Address: 200 W. Towsontown Boulevard, Baltimore, MD 21204-5200. Email: sfainfo@successforall.org. Website: http://www.successforall.net/Programs/readingwings.html. Telephone: (800) 548-4998 ext. 2372.

Program details

Cooperative Integrated Reading and Composition® was first used as part of a cooperative elementary whole-school reform model. The program was later reformulated as *Reading Roots* (for beginning readers) and *Reading Wings* (for upper elementary students) and is both a component of the *Success for All (SFA)* comprehensive school reform model and a stand-alone reading program.

The program uses daily 90-minute lessons to focus on story-related activities, direct instruction in reading comprehension, and integrated reading and language arts activities. In a team setting, mixed-ability students work together to read, discuss their reading to clarify unknown vocabulary, reread for fluency, understand the main idea, comprehend stories, and work through the writing process linked to the texts that the students are reading (including drafting, revising, and editing one another's writing). Students are rewarded on the basis of the team's performance to provide motivation to work together and help one another.

Teacher training includes a two-day session that covers word structure and phonics, vocabulary development, fluency, and comprehension skills, as well as program management and cooperative learning strategies. Technical support by phone or onsite visits also is provided.

Cost

The cost of the program is approximately $150 per student for training and materials, depending on school size and the number of schools within a district that are participating.

Research Summary

Thirty-eight studies reviewed by the WWC investigated the effects of *Cooperative Integrated Reading and Composition*® on beginning readers. One study (Stevens, Slavin, & Farnish, 1991) is a randomized controlled trial that meets WWC evidence standards without reservations. One study (Bramlett, 1994) is a quasi-experimental design that meets WWC evidence standards with reservations. These two studies are summarized in this report. The remaining 36 studies do not meet either WWC eligibility screens or evidence standards. (See references beginning on p. 6 for citations for all 38 studies.)

Table 2. Scope of reviewed research

Grade	3, 4
Delivery method	Small group/Whole class
Program type	Curriculum
Studies reviewed	38
Meets WWC standards without reservations	1 study
Meets WWC standards with reservations	1 study

Summary of study meeting WWC evidence standards without reservations

Stevens et al. (1991) examined the effects of *Cooperative Integrated Reading and Composition*® *(CIRC®)* in a cluster randomized trial of 30 classrooms, with 486 third- and fourth-grade students in four schools in Harrisburg, Pennsylvania. A total of 153 students in 10 intervention-group classrooms received *CIRC*®, and 167 students in 10 comparison-group classrooms received their regular reading curriculum.[4] Four days a week, students in the intervention classrooms spent half of their reading time using *CIRC*® materials. Teachers taught comprehension strategies and metacomprehension skills as a part of the *CIRC*® curriculum. Following this instruction, students worked in teams on follow-up activities. The classrooms in the comparison group used traditional methods and curriculum materials. These included the use of a basal reading series with related workbook and follow-up worksheet activities. The study reported student outcomes after four weeks of program implementation.

Summary of study meeting WWC evidence standards with reservations

Bramlett (1994) conducted a quasi-experiment of 18 classrooms (392 third graders) in eight school districts in rural southern Ohio. The reading components of *CIRC*® were implemented in the intervention classrooms as the core reading curriculum. The composition component of *CIRC*® was not used by the intervention classrooms participating in this study. The comparison classrooms received their regular reading curriculum. The study reported student outcomes after one school year of program implementation.

Effectiveness Summary

The WWC review of interventions for Beginning Reading addresses student outcomes in four domains: alphabetics, fluency, comprehension, and general reading achievement. The two studies that contribute to the effectiveness ratings in this report cover two domains: comprehension and general reading achievement. The findings below present the authors' estimates and WWC-calculated estimates of the size and the statistical significance of the effects of *Cooperative Integrated Reading and Composition*® on beginning readers. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 18.

Summary of effectiveness for the comprehension domain

Two studies reported findings in the comprehension domain.

Stevens et al. (1991) reported a statistically significant positive difference between the intervention group (which pooled the *CIRC*® and *Direct Instruction* intervention groups together), and the comparison group on the Main Idea Questions outcome. According to WWC calculations, the difference between the *CIRC*® group and the comparison group also was statistically significant. The authors did not find statistically significant differences between the pooled intervention group and the comparison group on the Inference Questions outcome, and the WWC calculations for the difference between the *CIRC*® and the comparison group confirmed the lack of a significant difference.

Bramlett (1994) reported a statistically significant positive difference between the *CIRC*® group and the comparison group on the Reading Comprehension subtest of the California Achievement Test (CAT). According to WWC calculations (which account for clustering and multiple comparisons), the difference was not statistically significant. The study author found, and the WWC confirmed, no statistically significant difference between the *CIRC*® group and the comparison group on the CAT Total Reading, CAT Word Analysis, and Reading Vocabulary subtests. (Note that the CAT Total Reading was comprised of Reading Vocabulary and Reading Comprehension.) The average effect size across the three outcomes was not large enough to be considered substantively important according to WWC criteria (i.e., larger than 0.25).

Thus, for the comprehension domain, one study shows a statistically significant positive effect and one study shows an indeterminate effect. This results in a domain rating of potentially positive 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 *Cooperative Integrated Reading and Composition*® in the comprehension domain had one study showing a statistically significant positive effect and one study showing an indeterminate effect.
Extent of evidence	**Criteria met**
Medium to large	The review of *Cooperative Integrated Reading and Composition*® in the comprehension domain was based on two studies that included more than 12 schools[5] and 712 students.

Summary of effectiveness for the general reading achievement domain

One study reported findings in the general reading achievement domain.

In a separate manuscript describing the same study as reported in Stevens et al. (1991), Stevens, Slavin, and Farnish (1989) provided results for the Iowa Test of Basic Skills (ITBS) Reading Achievement assessment, but they did not report the inferential results of a contrast between the *CIRC*® intervention group and the comparison group. The WWC calculations determined that there was not a statistically significant or substantively important difference between the groups on this outcome.

Thus, for the general reading achievement domain, one study shows an indeterminate effect. This results in a rating of no discernible effects, with a small extent of evidence.

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

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	The review of *Cooperative Integrated Reading and Composition*® in the general reading achievement domain had one study showing an indeterminate effect.
Extent of evidence	**Criteria met**
Small	The review of *Cooperative Integrated Reading and Composition*® in the general reading achievement domain was based on one study that included four schools and 320 students.

References

Study that meets WWC evidence standards without reservations

Stevens, R. J., Slavin, R. E., & Farnish, A. M. (1991). The effects of cooperative learning and direct instruction in reading comprehension strategies on main idea identification. *Journal of Educational Psychology, 83*(1), 8–16.

Additional source:

Stevens, R. J., Slavin, R. E., & Farnish, A. M. (1989). *The effects of cooperative learning and direct instruction in reading comprehension strategies on main idea identification.* Office of Educational Research and Improvement (ED), Washington, DC, report number 44.

Study that meets WWC evidence standards with reservations

Bramlett, R. K. (1994). Implementing cooperative learning: A field study evaluating issues for school-based consultants. *Journal of School Psychology, 32(*1), 67–84.

Studies that do not meet WWC evidence standards

Jenkins, J. R., Jewell, M., Leicester, N., O'Connor, R. E., Jenkins, L. M., & Troutner, N. M. (1994). Accommodations for individual differences without classroom ability groups: An experiment in school restructuring. *Exceptional Children, 60*(4), 344–358. 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.

Additional source:

Jenkins, J. R., Jewell, M., Leicester, N., Jenkins, L., & Troutner, N. M. (1991). Development of a school building model for educating students with handicaps and at-risk students in general education classrooms. *Journal of Learning Disabilities, 24*(5), 311–320.

Rapp, J. C. (1991). The effects of cooperative learning on selected student variables (Cooperative Integrated Reading and Composition on academic achievement in reading comprehension, vocabulary, and spelling on student self-esteem). *Dissertation Abstracts International, 52*(10), 3516A. The study does not meet WWC evidence standards because the measures of effect cannot be attributed solely to the intervention—there was only one unit assigned to one or both conditions.

Skeans, S. E. S. (1991). The effects of Cooperative Integrated Reading and Composition: Fidelity of implementation, and teacher concerns on student achievement. *Dissertation Abstracts International, 53*(02), 0455A. 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.

Studies that are ineligible for review using the Beginning Reading Evidence Review Protocol

Algozzine, B., Cooke, N., White, R., Helf, S., Algozzine, K., & McClanahan, T. (2008). The North Carolina Reading and Behavior Center's K–3 prevention model: Eastside Elementary School case study. In C. R. Greenwood, T. R. Kratochwill, & M. Clements (Eds.), *Schoolwide prevention models: Lessons learned in elementary schools* (pp. 173–214). New York: Guilford Press. 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.

Barr, A. (2006). *A descriptive study of reading strategies for secondary education in Minnesota Public Schools.* Menomonie: University of Wisconsin–Stout. 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.

Best Evidence Encyclopedia. (2009). *Reading/struggling readers.* Retrieved from http://www.bestevidence.org/reading/strug/strug_read.htm. 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.

Best Evidence Encyclopedia. (n.d.). *Cooperative Integrated Reading and Composition (CIRC) programs disseminated by the Success for All Foundation.* Retrieved from http://www.bestevidence.org/overviews/C/CIRC.pdf. 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.

Calderon, M., Hertz-Lazarowitz, R., & Slavin, R. E. (1998). Effects of bilingual Cooperative Integrated Reading and Composition on students making the transition from Spanish to English reading. *The Elementary School Journal, 99*(2), 153–165. This study is ineligible for review because it does not examine an intervention conducted in English.

Cooperative Integrated Reading and Composition. (2009). *Promising Practices Network.* Retrieved from http://www.promisingpractices.net/program.asp?programid=142. 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.

Croce, L. (2007). *The effect of cooperative learning on elementary school students' academic achievement and prosocial behavior.* Cortland, NY: State University College of Cortland. The study is ineligible for review because it does not use a comparison group or single-case design.

Durukan, E. (2011). Effects of Cooperative Integrated Reading and Composition (CIRC) technique on reading-writing skills. *Educational Research and Reviews, 6*(1), 102–109. 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.

Elmore, O. C. (2005). *Analysis of the principal's perceptions of the implementation and impact of the Accelerated Reader and other selected reading strategies used by Texas Gold Performance elementary schools.* College Station: Texas A&M University. The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.

Fashola, O. S., & Slavin, R. E. (1997). Promising programs for elementary and middle schools: Evidence of effectiveness and replicability. *Journal of Education for Students Placed at Risk, 2*(3), 251. 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.

Fuchs, D., & Fuchs, L. S. (2007). Increasing strategic reading comprehension with peer-assisted learning activities. In D. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 175–199). New York: Psychology Press. 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.

Jenkins, J. R., & O'Connor, R. E. (2005). Cooperative learning for students with learning disabilities: Evidence from experiments, observations, and interviews. In L. Swanson, S. Graham, & K. Harris (Eds.), *Handbook of learning disabilities* (pp. 417–430). New York: Guilford Press. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Lafont, L., Proeres, M., & Vallet, C. (2007). Cooperative group learning in a team game: Role of verbal exchanges among peers. *Social Psychology of Education, 10*(1), 93–113. The study is ineligible for review because it does not take place in the geographic area specified in the protocol.

Liao, H. C. (2005). *Effects of cooperative learning on motivation, learning strategy utilization, and grammar achievement of English language learners in Taiwan* (Unpublished doctoral dissertation). University of New Orleans, LA. This study is ineligible for review because it does not examine an intervention implemented in a way that falls within the scope of the review.

Meisinger, E. B., & Bradley, B. A. (2007). Classroom practices for supporting fluency development. In M. R. Kuhn & P. J. Schwanenflugel (Eds.), *Fluency in the classroom* (pp. 36–55). New York: Guilford Press. 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.

Nath, L. R. (1996). A peer tutoring training model for cooperative groupings: Is the effectiveness of cooperative groupings enhanced by students obtaining peer tutoring skills? *Dissertation Abstracts International, 57*(12), 5051A. The study is ineligible for review because it does not use a comparison group or single-case design.

Nelson, J. R., Lane, K. L., Benner, G. J., & Kim, O. (2011). A best evidence synthesis of literacy instruction on the social adjustment of students with or at-risk for behavior disorders. *Education & Treatment of Children, 34*(1), 141–162. 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.

Peurach, D. J. (2005). *Designing and managing comprehensive school reform: The case of Success for All* (Unpublished doctoral dissertation). University of Michigan, Ann Arbor. 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.

Ramey, S. L., & Ramey, C. T. (2006). Early educational interventions: Principles of effective and sustained benefits from targeted early education programs. In D. Dickinson & S. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 445–459). New York: Guilford Press. 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.

Roy, D. D. (2008). Assessing validity of web-based computer adaptive training modules. *Journal of the Indian Academy of Applied Psychology, 34*(1), 127–136. 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.

Scott, G. M., Lonergan, D. C., & Mumford, M. D. (2005). Conceptual combination: Alternative knowledge structures, alternative heuristics. *Creativity Research Journal, 17*(1), 79–98. 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.

Slavin, R. E. (2005). *Evidence-based reform: Advancing the education of students at risk.* Washington, DC: Center for American Progress and Institute for America's Future. 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., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research, 79*(4), 1391–1466. 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., Chambers, B., & Haxby, B. (2009). *2 million children: Success for All* (2nd ed.). Thousand Oaks, CA: Corwin Press. 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., & Datnow, A. (2007). Research in, research out: The role of research in the development and scale-up of Success for All. In S. Fuhrman, D. Cohen, & F. Mosher (Eds.), *The state of education policy research* (pp. 261–280). Mahwah, NJ: Lawrence Erlbaum Associates. 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.

Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative Integrated Reading and Composition: Two field experiments (Study: Spring 1985). *Reading Research Quarterly, 22*(4), 433–454. This study is ineligible for review because it does not provide enough information about its design to assess whether it meets standards.

Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative Integrated Reading and Composition: Two field experiments (Study: Fall 1985). *Reading Research Quarterly, 22*(4), 433–454. This study is ineligible for review because it does not provide enough information about its design to assess whether it meets standards.

Stevens, R. J., & Slavin, R. E. (1995). Effects of a cooperative learning approach in reading and writing on academically handicapped and nonhandicapped students. *The Elementary School Journal, 95*(3), 241–262. This study is ineligible for review because it does not disaggregate findings for the age or grade range specified in the protocol.

Stevens, R. J., & Slavin, R. E. (1995). The cooperative elementary school: Effects on students' achievement, attitudes, and social relations. *American Educational Research Journal, 32*(2), 321–351. This study is ineligible for review because it does not disaggregate findings for the age or grade range specified in the protocol

Thompson, J. C. (2005). *Cooperative learning in computer-supported classes.* Unpublished manuscript. 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.

Vaughn, S., Linan-Thompson, S., Pollard-Durodola, S. D., Mathes, P., & Cardenas-Hagan, E. (2005). Effective interventions for English language learners (Spanish/English) at-risk for reading difficulties. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 185–197). New York: Guilford Press. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Yeh, Y. (2007). *Elementary students' prior knowledge and the Cooperative Integrated Reading and Composition (CIRC) model in second-language reading comprehension* (Unpublished doctoral dissertation). Fordham University, New York. The study is ineligible for review because it does not examine an intervention conducted in English.

Appendix A.1: Research details for Stevens, Slavin, & Farnish, 1991

Stevens, R. J., Slavin, R. E., & Farnish, A. M. (1991). The effects of cooperative learning and direct instruction in reading comprehension strategies on main idea identification. *Journal of Educational Psychology, 83*(1), 8–16.

Table A1. Summary of findings

Meets WWC evidence standards without reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Comprehension	320 students	+20	Yes
General reading achievement	320 students	+1	No

Setting

This study took place in four schools in Harrisburg, Pennsylvania.

Study sample

The study is a classroom-level randomized controlled trial. A total of 30 third- and fourth-grade classrooms in four schools were randomly assigned to one of three conditions (balanced by grade):

1. 10 classrooms with 153 students were assigned to the *CIRC*® (direct instruction with cooperative learning) group,
2. 10 classrooms with 166 students were assigned to the direct instruction group without cooperative learning (DI), and
3. 10 classrooms with 167 students were assigned to the business-as-usual comparison group.

Both groups 1 and 2 used *CIRC*® materials on main idea comprehension during direct instruction, but group 2 did not use the cooperative learning component of *CIRC*®. Therefore, for the purposes of this report, the effects of *CIRC*® are estimated by comparing the *CIRC*® group against the business-as-usual comparison group. These results are shown in Appendices C.1 and C.2.

Intervention group

Four days each week, the direct instruction with cooperative learning group (group 1) spent half of its reading time using *CIRC*® materials. Teachers taught comprehension strategies and metacomprehension skills as a part of the *CIRC*® curriculum. Following this instruction, students worked in teams on follow-up activities. The study reported student outcomes after four weeks of program implementation.

Comparison group

The classrooms in the comparison group used traditional methods and curriculum materials. This included the use of a basal reading series with related workbook and follow-up worksheet activities.

Outcomes and measurement

Two investigator-developed assessments were used: One measured a student's ability to recall the main idea of a passage, and a second measured a student's ability to make correct inferences from a reading passage. End-of-year reading achievement scores from a standardized test, the Iowa Test of Basic Skills, also were used as outcomes in this study. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Teachers in the intervention condition received a one-day (six-hour) training in *CIRC*® by a certified trainer and received all of the supplemental materials necessary for the *CIRC*® reading program.

Appendix A.2: Research details for Bramlett, 1994

Bramlett, R. K. (1994). Implementing cooperative learning: A field study evaluating issues for school-based consultants. *Journal of School Psychology, 32*(1), 67–84.

Table A2. Summary of findings

Meets WWC evidence standards with reservations

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Comprehension	392 students	+4	No

Setting

The study took place in eight school districts in rural southern Ohio. The number of participating schools was not provided in the study.

Study sample

Eighteen third-grade teachers from eight school districts volunteered to participate in this quasi-experimental study. They were matched on school district and years of teaching experience and equally divided into two groups. In the analysis sample, the *CIRC*® group included 198 students in nine classrooms, and the comparison group included 194 students in nine classrooms. Each of the two groups of children was divided into three ability levels (lowest 33%, middle 33%, and upper 34%) based on the students' California Achievement Test (CAT) total reading score percentile rankings (administered prior to implementation of *CIRC*®). These subgroup results are presented in Appendix D.2.[6]

Intervention group

Students in the nine intervention classes were given only the reading components of the *CIRC*® program: basal-related activities, partner reading, story structure, words out loud, word meaning, story retelling, spelling, direct instruction in reading comprehension, and independent reading. The composition component of the *CIRC*® intervention was not used. The study reported student outcomes after one school year of program implementation.

Comparison group

Students in the comparison group received their regular reading curriculum, which was not described in the study. Teachers in the comparison group were promised *CIRC*® training at the completion of the study, and six of them were subsequently trained.

Outcomes and measurement

Teachers administered four CAT measures in the fall of 1990 and in the spring of 1991: Reading Vocabulary, Reading Comprehension, Total Reading, and Word Analysis. (Note that the Total Reading measure is comprised of Reading Vocabulary and Reading Comprehension.) Findings for the Total Reading and Word Analysis outcomes can be found in Appendix C.1. Subtest findings for Reading Vocabulary and Reading Comprehension can be found in Appendix D.1. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

The teachers received a one-day (six-hour) training in *CIRC*® by a certified trainer, as well as the project supplemental materials. Following training, the teachers were given assistance via observation and behavioral consultation sessions (approximately 15–30 minutes). Teachers also attended three half-day meetings during the study year to discuss implementation issues. The teachers in the comparison group were promised *CIRC*® training and materials upon completion of the study's collection of outcome data.

Appendix B: Outcome measures for each domain

Comprehension	
California Achievement Test (CAT): Total Reading (Form E)	This group-administered, standardized assessment is administered to grades 1 through 12 and consists of two subtests: Reading Comprehension and Reading Vocabulary. The Reading Comprehension subtest focuses on students' use of reading comprehension strategies. Passages reflect a wide range of narrative, expository, contemporary, and traditional texts. The subtest measures information recall, meaning construction, form analysis, and meaning evaluation of seven selections. The Reading Vocabulary subtest contains 20 items measuring same-meaning and opposite-meaning words, multi-meaning words, words in context, and the meaning of affixes (as cited in Bramlett, 1994).
Reading comprehension construct	
CAT: Word Analysis (Form E)	This is a group-administered, standardized assessment of word analysis. It is an optional subtest that is administered to students in grades K–3 and is not included as a component of the Total Battery score. This test measures a student's ability to recognize structural word parts, forms, vowels, consonants, and other phonetic forms (as cited in Bramlett, 1994).
Inference Questions	This is an author-developed, 10-item multiple-choice test that asks students to make inferences from each of 10 paragraphs (as cited in Stevens et al., 1991).
Main Idea Questions	This is an author-developed, 10-item multiple-choice test that asks students to identify the main idea of each of 10 paragraphs (as cited in Stevens et al., 1991).
Vocabulary development construct	
CAT: Reading Vocabulary (Form E)	This is a group-administered, standardized assessment of vocabulary. The Reading Vocabulary subtest contains 20 items measuring same-meaning and opposite-meaning words, multi-meaning words, words in context, and the meaning of affixes (as cited in Bramlett, 1994).
General reading achievement	
Iowa Test of Basic Skills	This is a group-administered, standardized assessment that measures students' general reading ability (as cited in Stevens et al., 1991).

Appendix C.1: Findings included in the rating for the comprehension domain by construct

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Bramlett, 1994[a]								
CAT: Total Reading	Grade 3	18 classes/ 392 students	687.0 (48.2)	682.0 (56.8)	4.0	0.08	+3	0.07
CAT: Word Analysis	Grade 3	18 classes/ 392 students	667.0 (43.3)	662.0 (49.7)	5.0	0.11	+4	0.07
Domain average for comprehension (Bramlett, 1994)						**0.09**	**+4**	**Not statistically significant**
Stevens et al., 1991[b]								
Inference Questions	Grades 3 and 4	20 classes/ 320 students	5.69 (2.14)	5.28 (2.08)	0.41	0.19	+8	> 0.05
Main Idea Questions	Grades 3 and 4	20 classes/ 320 students	6.40 (1.83)	4.74 (2.03)	1.66	0.85	+30	< 0.01
Domain average for comprehension (Stevens et al., 1991)						**0.52**	**+20**	**Statistically significant**
Domain average for comprehension across all studies						**0.31**	**+12**	**na**

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 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 each study's domain average was determined by the WWC. CAT= California Achievement Test. na = not applicable.

[a] For Bramlett (1994), a correction for clustering was needed but did not affect significance levels. The CAT: Reading Comprehension contrast was not found to be statistically significant, after adjusting for clustering and multiple comparisons. The *p*-values presented here were reported in the original study. All CAT outcomes were adjusted by pretest total reading scores on the CAT.

[b] For Stevens et al. (1991), a correction for multiple comparisons was needed but did not affect significance levels. The *p*-values presented here were reported in the original article for a difference between the comparison and the pooled intervention group (which consisted of the combined *CIRC*® and *Direct Instruction* intervention groups). The author-developed *Inference Questions* and *Main Idea Questions* were adjusted by each pretest score and by the Iowa Test of Basic Skills pretest scores.

Appendix C.2: Findings included in the rating for the general reading 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 et al., 1991[a]								
Iowa Test of Basic Skills (z-score)	Grades 3 and 4	20 classes/ 320 students	–0.07 (1.00)	–0.09 (1.02)	0.02	0.02	+1	> 0.05
Domain average for general reading achievement (Stevens et al., 1991)						**0.02**	**+1**	**Not statistically significant**
Domain average for general reading achievement across all studies						**0.02**	**+1**	**Not statistically significant**

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 calculated by the WWC.

[a] For Stevens et al. (1991), a correction for clustering was needed, and there were no author-reported significance levels for this test in the original study. The *p*-values presented here were calculated by the WWC. The group means presented here were adjusted for pretests. Pretest and posttest data for the *Iowa Test of Basic Skills* were presented in Stevens et al. (1989). The WWC calculated the intervention group mean using a difference-in-differences approach (see the WWC Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means.

Appendix D.1: Summary of 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	
Bramlett, 1994[a]								
CAT: Reading Comprehension	Grade 3	18 classes/ 392 students	687.0 (56.4)	681.0 (61.1)	6.0	0.10	+4	> 0.05
CAT: Reading Vocabulary	Grade 3	18 classes/ 392 students	684.0 (48.7)	682.0 (59.5)	2.0	0.04	+1	> 0.05

Table Notes: The supplemental findings presented in this table are additional subtest findings from the studies 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 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. CAT= California Achievement Test.

[a] For Bramlett (1994), corrections for clustering and multiple comparisons were needed and resulted in significance levels that differ from those in the original study. The *CAT: Reading Comprehension* contrast was not found to be statistically significant, after adjusting for clustering and multiple comparisons. The *p*-values presented here were reported in the original study. All CAT outcomes were adjusted by pretest total reading scores on the CAT.

Appendix D.2: Summary of subgroup 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	
Bramlett, 1994[a]								
CAT: Reading Comprehension	Grade 3/ medium ability	18 classes/ 151 students	698.0 (45.6)	695.0 (35.7)	3.0	0.07	+3	> 0.05
CAT: Reading Vocabulary	Grade 3/ medium ability	18 classes/ 151 students	694.0 (36.7)	693.0 (30.0)	1.0	0.03	+1	> 0.05
CAT: Total Reading	Grade 3/ medium ability	18 classes/ 151 students	697.0 (36.1)	694.0 (27.3)	3.0	0.09	+3	> 0.05
CAT: Word Analysis	Grade 3/ medium ability	18 classes/ 151 students	670.0 (29.9)	673.0 (38.3)	–3.0	–0.09	–3	> 0.05
CAT: Reading Comprehension	Grade 3/ high ability	18 classes/ 92 students	744.0 (32.7)	735.0 (35.5)	9.0	0.26	+10	> 0.05
CAT: Reading Vocabulary	Grade 3/ high ability	18 classes/ 92 students	736.0 (33.1)	738.0 (31.6)	–2.0	–0.06	–2	> 0.05
CAT: Total Reading	Grade 3/ high ability	18 classes/ 92 students	740.0 (25.8)	737.0 (28.2)	3.0	0.11	+4	> 0.05
CAT: Word Analysis	Grade 3/ high ability	18 classes/ 92 students	712.0 (38.2)	704.0 (37.1)	8.0	0.21	+8	> 0.05

Table Notes: The supplemental findings presented in this table are additional subgroup findings from the studies 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. CAT= California Achievement Test.

[a] For Bramlett (1994), a correction for clustering was needed but did not affect significance levels. The *p*-values presented here were reported in the original study. The *CIRC*® group means were adjusted for pretest. Pretest total reading scores on the CAT were used as a covariate. For Bramlett (1994), high-ability students are defined as the upper 34% of the sample, and medium-ability students are defined as the middle 33% of the sample.

Appendix D.3: Summary of alternate contrast 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	
Stevens et al., 1991[a]								
Contrast: Direct Instruction with Cooperative Learning (CIRC®) vs. Direct Instruction without Cooperative Learning								
Inference Questions	Grades 3 and 4	20 classes/ 319 students	5.69 (2.14)	5.60 (2.19)	0.09	0.04	+2	> 0.05
Main Idea Questions	Grades 3 and 4	20 classes/ 319 students	6.40 (1.83)	5.79 (1.89)	0.61	0.33	+13	0.09
Contrast: Direct Instruction without Cooperative Learning vs. comparison								
Inference Questions	Grades 3 and 4	20 classes/ 333 students	5.60 (2.19)	5.28 (2.08)	0.32	0.15	+6	> 0.05
Main Idea Questions	Grades 3 and 4	20 classes/ 333 students	5.79 (1.89)	4.74 (2.03)	1.05	0.53	+20	< 0.01

Table Notes: The supplemental findings presented in this table are additional alternate contrast findings from the studies 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.

[a] For Stevens et al. (1991), a correction for multiple comparisons was needed but did not affect significance levels. The *p*-values presented here were reported in the original study. For the second contrast, *Direct Instruction without Cooperative Learning vs. comparison,* the *p*-values are provided for a difference between the comparison and the two pooled treatments (*CIRC®* and *Direct Instruction*). The *CIRC®* group means were adjusted for pretest for the *Inference Questions* and *Main Idea Questions.* Iowa Test of Basic Skills pretest scores, as well as each pretest score, were used as covariates.

Appendix D.4: Summary of alternate contrast findings for the general reading 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 et al., 1991[a]								
Contrast: Direct Instruction with Cooperative Learning (CIRC®) vs. Direct Instruction without Cooperative Learning								
Iowa Test of Basic Skills (z-score)	Grades 3 and 4	20 classes/ 319 students	0.07 (1.00)	0.06 (0.97)	0.01	0.01	0	> 0.05
Contrast: Direct Instruction without Cooperative Learning vs. comparison								
Iowa Test of Basic Skills (z-score)	Grades 3 and 4	20 classes/ 333 students	–0.08 (0.97)	–0.09 (1.02)	0.01	0.01	0	> 0.05

Table Notes: The supplemental findings presented in this table are additional alternate contrast findings from the studies 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.

[a] For Stevens et al. (1991), corrections for clustering were needed for the *Iowa Test of Basic Skills* outcome, and there were no author-reported significance levels for this test in the original study. The *p*-values presented here were calculated by the WWC. The group means presented here were adjusted for pretests. Pretest and posttest results for the *Iowa Test of Basic Skills* were presented in Stevens et al. (1989). The WWC calculated the intervention group mean using a difference-in-differences approach (see the WWC Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means.

Endnotes

[1] The descriptive information for this program was obtained from the previous intervention report. The WWC requests that developers review the program description sections for accuracy from their perspective. The program description was provided to the developer in April 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] This report has been updated to include reviews of 27 studies that have been reviewed since July 2007. (The previous report was released in July 2007.) The 27 additional studies were not eligible for review under the Beginning Reading protocol. A complete list and disposition of all studies reviewed are provided in the references. The report includes reviews of all previous studies that met standards with reservations and resulted in a revised disposition of Skeans, 1991: 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. This revised disposition is due to a change in the review protocol. In particular, in the protocol version 1.0, a preintervention difference in baseline characteristics of 0.50 standard deviations or less along with statistical adjustment for baseline differences was sufficient to demonstrate equivalence in quasi-experimental studies. In the protocol version 2.1, if preintervention differences are 0.25 standard deviations or larger, then the study cannot meet standards (even after a statistical adjustment). The studies in this report were reviewed using WWC Evidence Standards, version 2.1, as described in the Beginning Reading review protocol version 2.1. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available

[3] For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 18. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies.

[4] A total of 166 students in 10 classrooms received a version of *Direct Instruction* that used *CIRC*® materials on main idea comprehension but did not include the cooperative learning component. Results from this group (*Direct Instruction*) are not included in the evidence rating for this report but are shown in Appendices D.3 and D.4.

[5] Bramlett (1994) does not report an exact number of participating schools.

[6] The study did not establish baseline equivalence of the intervention and comparison students in the lowest 33% subgroup; thus, the lowest subgroup is excluded from Appendix D.2.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, June). *Beginning Reading intervention report: Cooperative Integrated Reading and Composition*®. 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. 18.

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. 18.

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.