What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Beginning Reading July 16, 2007

Cooperative Integrated Reading and Composition®

Program description[1]

Cooperative Integrated Reading and Composition® (*CIRC*) is a comprehensive reading and writing program for students in grades 2 through 8. It includes story-related activities, direct instruction in reading comprehension, and integrated reading and language arts activities. Pairs of students (grouped either by or across ability levels) read to each other, predict how stories will end, summarize stories, write responses, 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 process. The *CIRC*® process includes teacher instruction, team practice, peer assessment, and team/partner recognition. A Spanish version of the program was also designed for grades 2–5.[2]

Research

Two studies of *CIRC*® met the WWC evidence standards with reservations. They included over 700 third-grade students in Ohio and Texas.[3] The WWC considers the extent of evidence for *CIRC*® to be moderate to large for comprehension. No studies that met WWC evidence standards with or without reservations addressed alphabetics, fluency, or general reading achievement.

Effectiveness

The *CIRC*® program was found to have no discernible effects for comprehension.

	Alphabetics	*Fluency*	*Comprehension*	*General reading achievement*
Rating of effectiveness	na	na	No discernible effects	na
Improvement index[4]	na	na	Average: + 4 percentile points Average: +1 to +8 percentile points	na

na = not applicable

1. The descriptive information for this program was obtained from publicly available sources: the research literature (Bramlett, 1994; Skeans, 1991). The WWC requests developers to review the program description sections for accuracy from their perspective. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review.
2. The What Works Clearinghouse (WWC) reviewed the effects of *Bilingual CIRC* on the reading achievement of English language learners and the findings are reported in a separate WWC intervention report.
3. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

(continued)

Additional program information[1]

Developer and contact

Developed by 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. Web: http://www.successforall.net/elementary/readingwings.htm. Telephone: (800) 548-4998 ext. 2372.

Scope of use

CIRC® was developed in 1983 by Bob Slavin and Nancy Madden. In 1985 *CIRC*® was used as part of a cooperative elementary whole-school model. The program was reformulated as *Reading Roots* (beginning readers) and *Reading Wings* (upper elementary), and is both a component of the Success for All comprehensive school reform model and a stand-alone reading program.

Teaching

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, pairs of mixed-ability students work together to read, discuss their reading to clarify unknown vocabulary, re-read for fluency, understand the main idea, comprehend stories, and work through the writing process (draft, revise, and edit each other's writing). Students are rewarded on the basis of the whole team's performance to provide motivation for peer work. Teams are rewarded based on the sum of team members' performance on these assessments, creating motivation for peer teaching.

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 through telephone conference or on-site visits is also provided.

Cost

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

Research

Eleven studies reviewed by the WWC investigated the effects of *CIRC*®. Two studies (Bramlett, 1994; Skeans, 1991) were quasi-experimental designs that met WWC evidence standards with reservations. The remaining nine studies did not meet WWC evidence screens.

Met evidence standards with reservations

Bramlett (1994) included 392 third-graders from 18 classrooms in eight school districts in rural southern Ohio. *CIRC*® was implemented in the intervention classrooms as the core reading curriculum.[5] The comparison classrooms received their regular reading curriculum.

Skeans (1991) is a study of the third-grade classrooms in a suburban district of Houston, Texas. Twenty-four third-grade teachers were matched on students' achievement and other factors and assigned to two conditions. In all, 169 students in the intervention group used *CIRC*® for 18 weeks along with the integrated language arts, and 141 students in the comparison group experienced only the integrated language arts curriculum.

4. These numbers show the average and range of improvement indices for all findings across the study.
5. Students in intervention classes were given only the reading components of *Cooperative Integrated Reading and Composition*®.

Research *(continued)*

Extent of evidence

The WWC categorizes the extent of evidence in each domain as small or moderate to large (see the What Works Clearinghouse Extent of Evidence Categorization Scheme). The extent of evidence takes into account the number of studies and the total sample size across the studies that met WWC evidence standards with or without reservations.[6] The WWC considers the extent of evidence for *CIRC*® to be moderate to large for comprehension. No studies that met WWC evidence standards with or without reservations addressed alphabetics, fluency, or general reading achievement.

Effectiveness

Findings

The WWC review of interventions for beginning reading addresses student outcomes in four domains: alphabetics, fluency, comprehension; and general reading achievement.[7] The studies included in this report covers only the comprehension domain. Within comprehension, results for two constructs, vocabulary development and reading comprehension, are reported. The findings below present the authors' estimates and WWC-calculated estimates of the size and the statistical significance of the effects of *CIRC*® on students.[8]

Comprehension. For reading comprehension, Bramlett (1994) reported a statistically significant positive effect of *CIRC*® on the Reading Comprehension subtest of the California Achievement Test (CAT). According to WWC calculations, the effect was not statistically significant. For vocabulary, the study authors did not find statistically significant effects of *CIRC*® on the CAT Word Analysis subtest nor on the CAT Reading Vocabulary subtest. The WWC found that the average effect size across the three comprehension outcomes was neither statistically significant nor large enough to be considered substantively important according to WWC criteria (that is, an effect size at least 0.25).

For reading comprehension, Skeans (1991) did not find a statistically significant effect of *CIRC*® on the Metropolitan Achievement Test (MAT) Reading Comprehension subtest. For vocabulary, the study author reported a statistically significant positive effect on the Vocabulary subtest of the MAT. According to WWC calculations, the effect was not statistically significant. The WWC found that the average effect size across the two outcomes was neither statistically significant not large enough to be substantively important.

Rating of effectiveness

The WWC rates the effects of an intervention in a given outcome domain as: positive, potentially positive, mixed, no discernible effects, potentially negative, or negative. The rating of effectiveness takes into account four factors: the quality of the research design, the statistical significance of the findings,[8] the size of the difference between participants in the intervention and the comparison conditions, and the consistency in findings across studies (see the WWC Intervention Rating Scheme).

6. The Extent of Evidence Categorization was developed to tell readers how much evidence was used to determine the intervention rating, focusing on the number and size of studies. Additional factors associated with a related concept, external validity, such as the students' demographics and the types of settings in which studies took place, are not taken into account for the categorization.
7. For definitions of the domains, see the Beginning Reading Protocol.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate the statistical significance. In the case of *CIRC*®, corrections for clustering and multiple comparisons were needed.

The WWC found *CIRC*® to have no discernible effects on comprehension

Improvement index

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

The average improvement index for comprehension is +4 percentile points in two studies, with a range of +1 to +8 percentile points across findings.

Summary

The WWC reviewed 11 studies on *CIRC*®. Two studies met WWC standards with reservations; the others did not meet WWC evidence screens. Based on the two studies, the WWC found no discernible effects in the comprehension domain. The evidence presented in this report may change as new research emerges.

References

Met 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.

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. (UMI No. 9217026)

Did not meet WWC evidence screens

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.[9]

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.[10]

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. (UMI No. 9717224)[11]

Rapp, J. C. (1991). The effect 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. (UMI No. 9207225)[10]

9. Intervention not relevant: the *BCIRC* intervention was designed to teach English to non-native speakers, which is not the focus of the Beginning Reading review.
10. Does not use a strong causal design: there was only one intervention and/or one comparison unit in each study condition, so the analysis could not separate the effects of the intervention from other factors.
11. The outcome measures are not shown to be valid or reliable: the outcome measures used in this study did not demonstrate adequate reliability or validity.

References *(continued)*

Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative integrated reading and composition: Two field experiments. *Reading Research Quarterly*, *22*(4), 433–454. **(Study: Fall 1985)**[12]

Stevens, R. J., Madden, N. A., Slavin, R. E., & Farnish, A. M. (1987). Cooperative Integrated Reading and Composition: Two field experiments. *Reading Research Quarterly, 22*(4), 433–454. **(Study: Spring 1985)**[12]

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.[13]

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.[13]

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.[13]

For more information about specific studies and WWC calculations, please see the WWC *CIRC*® Technical Appendices.

12. Complete data were not reported for the WWC to compute effect sizes for the third graders, the sample of interest to this review.
13. The sample is not appropriate to this review: the parameters for this WWC review specified that students should be in grades kindergarten through third grade; this study did not disaggregate students in the eligible range from those outside the range.

Appendix

Appendix A1.1 Study characteristics: Bramlett, 1994 (quasi-experimental design)

Characteristic	Description
Study citation	Bramlett, R. K. (1994). Implementing cooperative learning: A field study evaluating issues for school-based consultants. *Journal of School Psychology, 32*(1), 67–84.
Participants	Eighteen third-grade teachers volunteered to participate in the study. They were matched on the basis of geographic representativeness (school district) and years of teaching experience. 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 were divided into three ability levels (lowest 33%, middle 33%, and upper 34%) based on the students' percentile rankings of the pretest California Achievement Test (CAT) total reading scores.[1]
Setting	The study took place in eight school districts in rural southern Ohio. The number of participating schools was not provided in the study.
Intervention	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.
Comparison	Students in the comparison group received their regular reading curriculum, which was not indicated in the study. Teachers in the comparison group were promised that they could receive *CIRC*® training at the completion of the study, and six of them were subsequently trained.
Primary outcomes and measurement	The author administered four California Achievement Test (CAT) subtests: Reading Vocabulary, Reading Comprehension, Total Reading, and Word Analysis. The Total Reading subtest has not been included in this review because it is based on the results from the two subtests reported separately: Reading Vocabulary and Reading Comprehension (see Appendix A2 for more detailed descriptions of outcome measures).
Teacher training	The teachers received a one-day (6 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 training and materials upon completion of the first year of the project.

1. For the lowest 33% subgroup, the study did not establish that the comparison group was equivalent to the intervention group at baseline. Analyses of the other two subgroups (middle ability and higher ability) are presented in Appendix A4.

Appendix A1.2 Study characteristics: Skeans, 1991 (quasi-experimental design)

Characteristic	Description
Study citation	Skeans, Sharon 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. (UMI No. 9217026)
Participants	The study used intact classroom groups from the same school district, with twenty-four third-grade teachers matched on students' achievement, demographic factors, and hours of training.[1] No attrition was reported. In the analysis sample, 169 students were in the 12 intervention classrooms and 141 students were in the 12 comparison classrooms.
Setting	The study took place in a suburban district north of Houston, Texas.
Intervention	The intervention group was taught *CIRC*® integrated language arts and cooperative learning for 18 weeks. The program used daily 90-minute lessons to focus on story-related activities, direct instruction in reading comprehension, and integrated reading and language arts activities. Within four-member cooperative teams, pairs of mixed-ability students worked together to read. Within their teams, students re-read the story and discussed to clarify unknown vocabulary, understand the main idea, and comprehend stories. They worked through the writing process together, drafting, revising, and editing each other's writing.
Comparison	Students in the comparison classes continued using the district's integrated language arts program. The children's pretest scores on the Metropolitan Achievement Test were used as a covariant in the analyses to account for the differences between groups.
Primary outcomes and measurement	The author administered two subtests of the Metropolitan Achievement Test (MAT): Vocabulary and Reading Comprehension. Third graders received the elementary level of the test. The MAT Language subtest was also used in the study, but has not been included in this review because it was outside the scope of the Beginning Reading review (see Appendix A2 for more detailed descriptions of outcome measures).
Teacher training	The intervention teachers received at least 12 hours of *CIRC*® training. The teachers in both conditions had received at least 18 hours of training in the district's integrated language arts program. In order to continue teaching with the program, teachers needed to return a Stages of Concern questionnaire and a formative evaluation of *CIRC*®, indicating that they were implementing the program.

1. The fifth-grade sample that was included in this study is not reviewed in this report because it is outside the scope of the review. For sample relevancy criteria please see the Beginning Reading Protocol.

Appendix A2 Outcome measures in the comprehension domain by construct

Outcome measure	Description
Reading comprehension	
California Achievement Test (CAT): Reading Comprehension (Form E)	A group-administered, standardized assessment of reading comprehension (as cited in Bramlett, 1994).
CAT: Word Analysis (Form E)	A group-administered, standardized assessment of word analysis (as cited in Bramlett, 1994).
Metropolitan Achievement Test-6 (MAT-6): Reading Comprehension (Form L)	A group-administered, standardized assessment of reading comprehension (as cited in Skeans, 1991).
Vocabulary development	
CAT: Reading Vocabulary (Form E)	A group-administered, standardized assessment of vocabulary (as cited in Bramlett, 1994).
MAT-6: Vocabulary subtest (Form L)	A group-administered, standardized assessment of vocabulary (as cited in Skeans, 1991).

Appendix A3 Summary of study findings included in the rating for the comprehension domain by construct[1]

Outcome measure	Study sample	Sample size (classrooms/ students)	Authors' findings from the study: Mean outcome (standard deviation[2]) *CIRC*® group[3]	Authors' findings from the study: Mean outcome (standard deviation[2]) Comparison group	WWC calculations: Mean difference[4] (*CIRC*® – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Construct: Reading comprehension								
Bramlett, 1994 (quasi-experimental design)[8]								
CAT: Reading Comprehension	Third grade	18/392	687.0 (56.4)	681.0 (61.1)	6.0	0.10	ns	+4
CAT: Word Analysis	Third grade	18/392	667.0 (43.3)	662.0 (49.7)	5.0	0.11	ns	+4
Skeans, 1991 (quasi-experimental design)[8]								
MAT-6: Reading Comprehension	Third grade	24/305	58.6 (20.7)	57.0 (20.6)	1.6	0.08	ns	+3
Construct: Vocabulary development								
Bramlett, 1994 (quasi-experimental design)[8]								
CAT: Reading Vocabulary	Third grade	18/392	684.0 (48.7)	682.0 (59.5)	2.0	0.04	ns	+1
Skeans, 1991 (quasi-experimental design)[8]								
MAT-6: Vocabulary	Third grade	24/310	61.2 (17.9)	57.5 (18.3)	3.7	0.20	ns	+8
Average[9] for comprehension domain (Bramlett, 1994)						0.08	ns	+3
Average[9] for comprehension domain (Skeans, 1991)						0.14	ns	+6
Domain average[9] for comprehension domain						0.11	na	+4

ns = not statistically significant
na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the improvement index. Subgroup findings for high ability students (defined as upper 34% of the sample) and medium ability students (middle 33%) are presented in Appendix A4.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The *CIRC*® group means were adjusted for pretest. Pretest reading scores were used as a covariant.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.

(continued)

Appendix A3 Summary of study findings included in the rating for the comprehension domain by construct[1] *(continued)*

6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition versus the percentile rank of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Bramlett (1994), a correction for clustering and multiple comparisons was needed, so the significance levels may differ from those reported in the original study. In the case of Skeans (1991), a correction for clustering was needed, so the significance levels may differ from those reported in the original study.
9. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect size.

Appendix A4 Summary of subgroup findings for the comprehension domain by construct[1]

Outcome measure	Study sample	Sample size (classrooms/ students)	Authors' findings from the study: Mean outcome (standard deviation[2]) — *CIRC*® group[3]	Authors' findings from the study: Mean outcome (standard deviation[2]) — Comparison group	WWC calculations: Mean difference[4] (*CIRC*® – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Construct: Reading comprehension								
Bramlett, 1994 (quasi-experimental design)[8]								
CAT: Reading Comprehension	Third grade/ medium	18/151	698.0 (45.6)	695.0 (35.7)	3.0	0.07	ns	+3
CAT: Word Analysis	Third grade/ medium	18/151	670.0 (29.9)	673.0 (38.3)	–3.0	–0.09	ns	–3
CAT: Reading Comprehension	Third grade/ high	18/92	744.0 (32.7)	735.0 (35.5)	9.0	0.26	ns	+10
CAT: Word Analysis	Third grade/ high	18/92	712.0 (38.2)	704.0 (37.1)	8.0	0.21	ns	+8
Construct: Vocabulary development								
Bramlett, 1994 (quasi-experimental design)[8]								
CAT: Reading Vocabulary	Third grade/ medium	18/151	694.0 (36.7)	693.0 (30.0)	1.0	0.03	ns	+1
CAT: Reading Vocabulary	Third grade/ high	18/92	736.0 (33.1)	738.0 (31.6)	–2.0	–0.06	ns	–2

ns = not statistically significant

1. This appendix presents subgroup findings for high ability students (defined as upper 34% of the sample) and medium ability students (middle 33%) for measures that fall in the comprehension domain. Total group scores were used for rating purposes and are presented in Appendix A3.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The *CIRC*® group means were adjusted for pretest. Pretest reading scores were used as a covariant.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition versus the percentile rank of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Bramlett (1994), a correction for clustering was needed, so the significance levels may differ from those reported in the original study.

Appendix A5 *CIRC*® rating for the comprehension domain

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

For the outcome domain of comprehension, the WWC rated *CIRC*® as having no discernible effects. It did not meet the criteria for other ratings (positive effects, potentially positive effects, mixed effects, potentially negative effects, and negative effects) because the one study that met WWC standards did not show statistically significant or substantively important effects.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative.*

  Met. No study showed a statistically significant or substantively important effect, either positive or negative.

Other ratings considered

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

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

  Not met. No study showed statistically significant positive effects or met the WWC evidence standards for a strong design.

AND

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

  Met. No study showed statistically significant or substantively important negative effects.

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

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

  Not met. No study showed a statistically significant or substantively important positive effect.

AND

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

  Not met. No study showed a statistically significant or substantively important negative effect, but two studies showed indeterminate effects.

Mixed effects: Evidence of inconsistent effects as demonstrated through either of the following criteria.

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

  Not met. No study showed a statistically significant or substantively important effect, either positive or negative.

OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

  Not met. No study showed a statistically significant or substantively important effect, while two studies showed indeterminate effects.

Appendix A5 *CIRC*® rating for the comprehension domain *(continued)*

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

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

  Not met. No study showed a statistically significant or substantively important negative effect.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

  Met. No study showed a statistically significant or substantively important positive effect.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

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

  Not met. No study showed a statistically significant or substantively important negative effect.

AND

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

  Met. No study showed statistically significant or substantively important positive effects.

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

Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence[2]
		Schools[1]	Students	
Alphabetics	0	0	0	na
Fluency	0	0	0	na
Comprehension	2	over 8	702	Moderate to large
General reading achievement	0	0	0	na

na = not applicable/not studied

1. No information is provided about the number of schools. Bramlett (1994) study took place in eight school districts (and 18 classrooms), while 24 teachers participated in Skeans (1991) study.
2. A rating of "moderate to large" requires at least two studies and two schools across studies in one domain and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is "small."