What Works Clearinghouse

ies INSTITUTE OF EDUCATION SCIENCES

Corrective Reading

Program Description[1]

Corrective Reading is designed to promote reading accuracy (decoding), fluency, and comprehension skills of students in grades 4–12 who are reading below their grade level. The program includes four sequential levels that address students' decoding skills and six sequential levels that address students' comprehension skills. The levels are designed to target students who need assistance with particular types of reading skills based on the results of *Corrective Reading* placement tests (see the Teaching section for more details on the different levels). The decoding and comprehension components can be used separately as a supplemental reading intervention or combined for use as a reading intervention curriculum. All lessons in the program are sequenced and scripted. *Corrective Reading* can be implemented in small groups of four to five students or in a whole-class format. *Corrective Reading* is intended to be taught in 45-minute lessons four to five times a week.

Research[2]

One study of *Corrective Reading* that falls within the scope of the Adolescent Literacy review protocol meets What Works Clearinghouse (WWC) evidence standards, and no studies meet WWC evidence standards with reservations. This study included 86 fifth-grade struggling readers from a school district just outside Pittsburgh, PA.[3]

Based on one study, the WWC considers the extent of evidence for *Corrective Reading* on adolescent learners to be small for alphabetics, reading fluency, and comprehension. The one study that meets WWC evidence standards did not examine the effectiveness of *Corrective Reading* on adolescent learners in the alphabetic, reading fluency, or general literacy achievement domains.

1. The descriptive information for this program was obtained from a publicly available source: the program's website (http://www.sraonline.com/, downloaded December 2009). 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. The literature search reflects documents publicly available by April 2009.
2. The studies in this report were reviewed using WWC Evidence Standards, Version 2.0 (see the WWC Procedures and Standards Handbook, Chapter III), as described in protocol Version 2.0.
3. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

Effectiveness

Corrective Reading was found to have no discernible effects on the alphabetics, reading fluency, and comprehension domains for adolescent learners.

	Alphabetics	*Reading fluency*	*Comprehension*	*General literacy achievement*
Rating of effectiveness	No discernible effects	No discernible effects	No discernible effects	na
Improvement index[4]	Average: +4 percentile points	+4 percentile points	Average: +3 percentile points	na
	Range: +1 to +6 percentile points	na	Range: +1 to +5 percentile points	na

na = not applicable

Absence of conflict of interest

The study in this intervention report, Torgesen et al. (2006), was prepared, in part, by staff of Mathematica Policy Research. For this reason, the study was rated by, and this intervention report was prepared by, researchers unaffiliated with Mathematica. The report was then reviewed by the principal investigator, a WWC Quality Assurance reviewer, and an external peer reviewer.

Additional program information

Developer and contact

Corrective Reading is distributed by SRA/McGraw-Hill. Address: 220 East Danieldale Road, Desoto, TX 75115-2490. Web: http://www.sraonline.com/. Telephone: (888) 772-4543.

Scope of use

Corrective Reading has been implemented in the United States and England. No information is available on the number of students or schools using the program.

Teaching

The program's 45-minute lessons are designed for groups of up to 20 students up to five times a week. The program's two components—decoding and comprehension—have four and six sequential levels of difficulty, respectively. Students' skill development is designed to progress as they move from lower to higher levels. For example, a student who needs assistance developing basic decoding skills would start at decoding level A and complete that level before moving on to the more advanced skills covered in level B1, whereas a student who does not read fluently or who confuses similar words would start at decoding level B1 and complete that level before moving on to the more advanced skills covered in level C.

Each level spans half of an academic year (with the exception of level C, which spans an entire academic year, and Fast Cycle levels that span roughly a quarter of a year). Therefore, the number of levels a student covers in a single academic year can range from one to three. The program can be used to provide students with either two full periods of instruction per day—one period in decoding and one period in comprehension—or one period of instruction per day (by focusing on one of the two components [e.g., comprehension]).

Corrective Reading placement tests determine the level at which each student is placed; once placed at a particular level, the program calls for the student to complete all the lessons in that level before moving on to the next level. All levels contain ongoing mastery tests and assessments to help track individual student achievement.

4. These numbers show the average and range of student-level improvement indices for all findings across the study.

Additional program information *(continued)*

The decoding levels include

- level A (65 lessons), which is designed for nonreaders. This level emphasizes basic decoding skills: rhyming, sounding out, sentence reading, and story reading;
- levels B1 and B2 (65 lessons in each), which are designed for struggling readers who do not read fluently or who confuse similar words. These levels teach students to become automatic decoders, able to read 90 words per minute by the end of B1 and 130 words per minute by the end of B2; and
- level C (125 lessons), which is designed for students who experience difficulty with vocabulary and complex sentence structures. This level bridges the gap between advanced word decoding skills and the ability to read informational text.

The comprehension levels include

- level A (65 lessons), which is designed for students who do not understand the concepts underlying much of the material being taught in classrooms;
- levels B1 and B2 (60 and 65 lessons, respectively), which target more advanced readers;[5] and
- level C (140 lessons), which focuses on applying comprehension skills.

The development of skills in the comprehension component progresses from comprehending oral language to comprehending written material. Skills are first taught in structured exercises that are controlled by the teacher. Later, students are shown how to apply the skills independently to complex written materials.

The publisher also provides staff development training that focuses on how to deliver direct instruction and use the program materials. Follow-up observations and coaching are recommended as support for teachers implementing the program. A Teaching Tutor CD-ROM provides ongoing support for teachers using *Corrective Reading*.

Cost

Prices vary by level (A, B1, B2, C) and component (decoding, comprehension). The cost of student materials ranges from $10 per student for level A programs to $50 per student for level C materials. Teacher materials cost approximately $200 per level.

For more detailed cost information by level and component, consult the distributor's website: http://www.sraonline.com/.

Research

A total of 129 studies reviewed by the WWC investigated the effects of *Corrective Reading* on adolescent learners. One study (Torgesen et al., 2006) is a randomized controlled trial that meets WWC evidence standards. The remaining 128 studies do not meet either WWC evidence standards or eligibility screens.

Meets evidence standards

Torgesen et al. (2006) conducted a randomized controlled trial that examined the effects of the decoding component of *Corrective Reading* on 86 fifth-grade students in Pennsylvania. The study design was based on random assignment of 32 school units[6] to one of four interventions: *Corrective Reading, Kaplan SpellRead*,[7] *Failure Free Reading*, and *Wilson Reading*. Within each school, eligible students were randomly assigned to the treatment group that would receive the intervention assigned to its school or to the control group that would not receive any of the four interventions. Students were eligible for participation if their teacher identified them as a struggling reader and if they scored at or below the 30th percentile on a word-level reading test and at or above the 5th percentile on a vocabulary test. The

5. Comprehension levels A and B1 also offer Fast Cycle alternatives that contain 30 and 35 lessons, respectively.
6. A school unit consists of several schools partnering so that the cluster included two 3rd-grade and two 5th-grade instructional groups. Only the findings on 5th graders are included in this review as specified by the Adolescent Literacy review protocol.
7. The study's authors refer to the intervention as *SpellRead P.A.T.* In 2006, Kaplan K12 acquired SpellRead, the developer and distributor of *SpellRead Phonological Auditory Training®*.

Research *(continued)*

WWC based its effectiveness ratings on findings from comparisons of the 55 fifth-grade students who received *Corrective Reading* and the 31 fifth-grade control group students who received the standard district curriculum. The study reported student outcomes after six months of program implementation.[8]

Extent of evidence

The WWC categorizes the extent of evidence in each domain as small or medium to large (see the WWC Procedures and Standards Handbook, Appendix G). The extent of evidence takes into account the number of studies and the total sample size across the studies that meet WWC evidence standards with or without reservations.[9]

The WWC considers the extent of evidence for *Corrective Reading* to be small for the alphabetics, reading fluency, and comprehension domains for adolescent learners. The one study that meets WWC evidence standards did not examine the effectiveness of *Corrective Reading* on adolescent learners in the alphabetic, reading fluency, or general literacy achievement domains.

Effectiveness

Findings

The WWC review of interventions for Adolescent Literacy addresses student outcomes in four domains: alphabetics, reading fluency, comprehension, and general literacy achievement. The study included in this report covers three domains: alphabetics, reading fluency, and comprehension. The findings below present the authors' estimates and WWC-calculated estimates of the size and the statistical significance of the effects of *Corrective Reading* on adolescent learners.[10]

Alphabetics. Torgesen et al. (2006) did not find statistically significant effects of *Corrective Reading* on 5th graders' scores on the Word Attack and Word Identification subtests of the Woodcock Reading Mastery Test–Revised (WRMT-R) or the Phonemic Decoding Efficiency and Sight Word Efficiency subtests of the Test of Word Reading Efficiency (TOWRE). The WWC-calculated average effect across these measures was not large enough to be considered substantively important according to WWC criteria (i.e., an effect size of at least 0.25).

Reading fluency. Torgesen et al. (2006) did not find statistically significant effects of *Corrective Reading* on 5th graders' scores on the Oral Reading Fluency test. The WWC-calculated effect was not large enough to be considered substantively important according to WWC criteria.

Comprehension. Torgesen et al. (2006) examined two outcomes in this domain (the WRMT-R Passage Comprehension subtest and the Group Reading Assessment and Diagnostic Evaluation [GRADE] Passage Comprehension subtest) and reported no statistically significant effects for 5th-grade students, although the comprehension component of *Corrective Reading* was not implemented in this study. The WWC-calculated average effect across these measures was not large enough to be considered substantively important according to WWC criteria.

8. For the purposes of this study, only the decoding component of *Corrective Reading* was implemented. By design (to facilitate the examination of two types of interventions), the comprehension component of *Corrective Reading* was not implemented. Additional findings reflecting students' outcomes one year after the intervention year can be found in Appendices A4.1–A4.3.
9. 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. Information about how the extent of evidence rating was determined for *Corrective Reading* is in Appendix A6.
10. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Torgesen et al. (2006), the authors adjusted for clustering and no corrections for multiple comparisons were needed because there were no statistically significant findings.

Effectiveness *(continued)*

In summary, the study showed indeterminate effects in the alphabetics, reading fluency, and comprehension domains.

Rating of effectiveness

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

The WWC found *Corrective Reading* to have no discernible effects on alphabetics, reading fluency, or comprehension for adolescent learners

Improvement index

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

The average improvement index for alphabetics is +4 percentile points, with a range of +1 to +6 percentile points across findings from one study. The improvement index for reading fluency is +4 percentile points for a single finding from one study. The average improvement index for comprehension is +3 percentile points, with a range of +1 to +5 percentile points across findings from one study.

Summary

The WWC reviewed 129 studies on *Corrective Reading* for adolescent learners. One of these studies meets WWC evidence standards; the remaining 128 studies do not meet either WWC evidence standards or eligibility screens. Based on one study, the WWC found no discernible effects on alphabetics, reading fluency, or comprehension for adolescent learners. The conclusions presented in this report may change as new research emerges.

References

Meets WWC evidence standards

Torgesen, J., Myers, D., Schirm, A., Stuart, E., Vartivarian, S., Mansfield, W., et al. (2006). *National assessment of Title I. Interim report. Volume II: Closing the reading gap: First year findings from a randomized trial of four reading interventions for striving readers*. Washington, DC: National Center for Education Evaluation and Regional Assistance.

Additional source:

Torgesen, J., Schirm, A., Castner, L., Vartivarian, S., Mansfield, W., Myers, D., et al. (2007). *National assessment of Title I. Final report. Volume II: Closing the reading gap: Findings from a randomized trial of four reading interventions for striving readers* (NCEE 2008-4013). Washington, DC: National Center for Education Evaluation and Regional Assistance.

Studies that fall outside the Adolescent Literacy review protocol or do not meet WWC evidence standards

Airhart, K. M. (2005). The effectiveness of Direct Instruction in reading compared to a state-mandated language arts curriculum for ninth and tenth graders with specific learning disabilities (Doctoral dissertation, Tennessee State University, 2005). *Dissertation Abstracts International, 67*(01A), 138–130.

References *(continued)*

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.

Alliance for Excellent Education. (n.d.). *Case study: Charles M. Goethe Middle School* (Corrective Reading model). Washington, DC: Author. The study is ineligible for review because it does not use a comparison group.

Anderson, L. L., & Alber, S. R. (2003). Precision teaching in a day treatment facility. *Journal of Precision Teaching & Celeration, 19*(1), 35–37. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Arthur, C. (1988). Progress in a high school LD class. *ADI News, 27*(4), 17–18. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Additional source:

SRA/McGraw–Hill. (n.d.). *Individuals with learning disabilities*. Desoto, TX: Author.

Baumann, J. F. (1986). *Teaching main idea comprehension*. Newark, DE: International Reading Association. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Benner, G. J., Kinder, D., Beaudoin, K. M., & Stein, M. (2005). The effects of the *Corrective Reading* decoding program on the basic reading skills and social adjustment of students with high-incidence disabilities. *Journal of Direct Instruction, 5*(1), 67. 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.

Bergeson, T. (2004). *Grades 4–12 reading intervention materials review: Washington State evaluation report*. Olympia, WA: Washington State Department of Education. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Betesh, S. (2007). *The effectiveness of the SRA* Corrective Reading for Comprehension *program for students with autism*. Unpublished master's thesis, Gwynedd-Mercy College, Gwynedd Valley, PA. 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.

Bradford, S., Alberto, P., Houchins, D. E., Shippen, M. E., & Flores, M. (2006). Using systematic instruction to teach decoding skills to middle school students with moderate intellectual disabilities. *Education & Training in Developmental Disabilities, 41*(4), 333–343. The study is ineligible for review because it does not use a comparison group.

Brosh, E. (2008). *Examining the impact of direct instruction on the reading, comprehension, and mathematic skills of adolescents with brain injuries*. Unpublished master's thesis, Southern Illinois University, Carbondale. The study is ineligible for review because it does not use a comparison group.

Brown, A. L. (2003). *Effectiveness of SRA* Corrective Reading. Unpublished master's thesis, California State University–Stanislaus. The study is ineligible for review because it does not use a comparison group.

Brynildson, K. D. S. (1990). The effects of audiotape self-monitoring on *Direct Instruction* teacher presentation techniques (Master's thesis, Western Michigan University, 1990). *Masters Abstracts International, 29*(01), 50–13. The study is ineligible for review because it does not include a student outcome.

Byron, D. (1988). *Corrective Reading* in a comprehensive school: The Hartcliffe project. *Educational and Child Psychology, 5*(4), 35–41. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Campbell, M. (1988). *Corrective Reading* program evaluated with secondary students in San Diego. *Direct Instruction News, 7*(4), 15–17. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Campbell, M. L. (1983). *A study of* Corrective Reading *as an effective and appropriate program for reading disabled, learning handicapped secondary students*. Unpublished report presented to the Faculty of the School of Education, San Diego State University. The study is ineligible for review

because it does not occur within the time frame specified in the protocol.

Campbell, M. L. (1984). *Corrective Reading* program evaluated with secondary students in San Diego. *ADI News, 3*(3). The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Canataro, J. F. (1988). A comparison of *Corrective Reading* and basal reader programs at the sixth grade level in urban schools (Doctoral dissertation, Temple University, 1988). *Dissertation Abstracts International, 49*(07A), 92–1746. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Cheung, A., & Slavin, R. E. (2005). Effective reading programs for English language learners and other language-minority students. *Bilingual Research Journal, 29*(2), 241. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Cihon, T. M., Gardner, R., Morrison, D., & Paul, P. V. (2008). Using visual phonics as a strategic intervention to increase literacy behaviors for kindergarten participants at-risk for reading failure. *Journal of Early & Intensive Behavior Intervention, 5*(3), 138–155. 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.

Clark, D. S. (2001). Components of effective reading instruction for reading disabled students: An evaluation of a program combining code- and strategy-instruction (Doctoral dissertation, University of Toronto, Canada, 2001). *Dissertation Abstracts International*, *62*(11A), 129–3727. 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.

Clunies-Ross, G. (1990). Some effects of Direct Instruction in comprehension skills with sixth grade students. *Behaviour Change, 7*(2), 84–89. 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.

Crane, L. R., Gustafson, J., & Poziemski, C. (2000). Motivational aspects of reading and its measurement in community college students. *Journal of Applied Research in the Community College, 7*(2), 87. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Davenport, J. L. (2004). *Teacher efficacy and Direct Instruction in reading*. Unpublished doctoral dissertation, University of Georgia, Athens. The study is ineligible for review because it does not use a comparison group.

Deshler, D. D., Palincsar, A. S., Biancarosa, G., & Nair, M. (2007). *Informed choices for struggling adolescent readers: A research-based guide to instructional programs and practices*. New York: Carnegie Corporation of New York. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Diliberto, J. A. (2006). Effects of a supplemental syllable skills curriculum on reading and spelling achievement in middle school students taught in remediation classes (Doctoral dissertation, University of North Carolina at Charlotte, 2006). *Dissertation Abstracts International, 67*(02A), 130–516. The study is ineligible for review because it does not use a comparison group.

Dolezal, D. N., Weber, K. P., Evavold, J. J., Wylie, J., & McLaughlin, T. F. (2007). The effects of a reinforcement package for on-task and reading behavior with at-risk and middle school students with disabilities. *Child & Family Behavior Therapy, 29*(2), 9–25. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Drakeford, W. (2002). The impact of an intensive program to increase the literacy skills of incarcerated youth. *Journal of Correctional Education, 53*(4), 139–144. The study is ineligible for review because it does not examine an intervention implemented in a way that falls within the scope of the review.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis

of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, *79*(1), 262–300. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Edmondson, C. A. (2004). Effects of instruction in the *Corrective Reading* and Voyager Reading programs on the reading skills of students enrolled in a university program (Doctoral dissertation, Utah State University, 2004). *Dissertation Abstracts International*, *65*(11A), 158–4147. 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.

Esham, L. A. (2001). *Effects of teacher training and coaching on student achievement in the* Corrective Reading *program at Sussex Central Middle School*. Unpublished doctoral dissertation, University of Delaware, Newark. The study is ineligible for review because it does not use a comparison group.

Fabrizio, M. A., Schirmer, K., & Ferris, K. (2002). Tracking curricular progress with precision. *Journal of Precision Teaching & Celeration, 18*(2), 78–79. The study is ineligible for review because it does not use a comparison group.

Figueredo, A. (1989). *A study to determine the effectiveness of* Corrective Reading *programs on nineteen students identified as mildly handicapped*. Unpublished master's thesis, Boise State University, ID. The study is ineligible for review because it does not use a comparison group.

Flores, M. M., & Ganz, J. B. (2007). Effectiveness of Direct Instruction for teaching statement inference, use of facts, and analogies to students with developmental disabilities and reading delays. *Focus on Autism and Other Developmental Disabilities, 22*(4), 244–251. 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.

Florida Center for Reading Research. (2004). *What is* Corrective Reading? Tallahassee, FL: Author. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Froelich, K. S. (1992). Comparing a literature-based and skills-based approach for elementary pupils of limited reading proficiency (Doctoral dissertation, Fordham University, 1992). *Dissertation Abstracts International*, *53*(09A), 157–3156. 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.

Gersten, R., Brockway, A., & Henares, N. (1983). The Monterey DI program for students with limited English (ESL). *Direct Instruction News, 2*(4), 8–9. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Additional source:

SRA/McGraw-Hill. (n.d.). *Limited English speakers*. Desoto, TX: Author.

Gibbs, C. (2004). *Reading decoding and fluency of high school students with special needs: The effectiveness of* Corrective Reading. Unpublished master's thesis, Cardinal Stritch University, Milwaukee, WI. The study is ineligible for review because it does not use a comparison group.

Glang, A., Singer, G., Cooley, E., & Tish, N. (1991). Direct Instruction: Applications with students with brain injury. *ADI News, 11*(1), 23–28. 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.

Additional source:

SRA/McGraw-Hill. (n.d.) *Individuals with brain injuries*. Desoto, TX: Author.

Goodwin, A. D. (2006). *Effectiveness of* Corrective Reading *with high school special education students*. Unpublished doctoral dissertation, University of Wisconsin, Eau Claire. 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.

Gregory, R., Hackney, C., & Gregory, N. M. (1982). *Corrective Reading* programme: An evaluation. *British Journal of Educational Psychology, 52*, 33–50. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Additional source:

SRA/McGraw-Hill. (n.d.). *Remedial readers, England*. Desoto, TX: Author.

Grossen, B. (2002). *Direct Instruction model for secondary schools: The research base for the* REACH *system*. New York: Center for Applied Research in Education. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Grossen, B. (2004). Success of a Direct Instruction model at a secondary level school with high-risk students. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 20*(2), 161–178. 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.

Grossen, B. (2008). *The research base for* Corrective Reading *SRA*. Eugene, OR: University of Oregon. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Guerrero, F. (1987). *Chapter I clinical and guidance program 1985–86: OEA evaluation report*. New York: New York City Board of Education, Office of Educational Assessment. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Gunn, B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *The Journal of Special Education, 34*(2), 90–103. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Additional sources:

Gunn, B., Smolkowski, K., Biglan, A., & Black, C. (2002). Supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school: A follow-up study. *The Journal of Special Education, 36*(2), 69–79.

Gunn, B., Smolkowski, K., Biglan, A., Black, C., & Blair, J. (2005). Fostering the development of reading skill through supplemental instruction: Results for Hispanic and non-Hispanic students. *Journal of Special Education, 39*(2), 66–85.

Harris, R. E., Marchand-Martella, N., & Martella, R. C. (2000). Effects of a peer-delivered *Corrective Reading* program. *Journal of Behavioral Education, 10*(1), 21–36. The study is ineligible for review because it does not use a comparison group.

Hempenstall, K. (2008). *Corrective Reading*: An evidence-based remedial reading intervention. *Australasian Journal of Special Education, 32*(1), 23–54. 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.

Herr, C. M. (1989). Using *Corrective Reading* with adults. *ADI News, 8*(2), 18–21. The study is ineligible for review because it does not use a comparison group.

Hicks, R. M. (1990). *A learning activity for at-risk third grade students in visual discrimination of the consonants B, D, and P*. Unpublished master's practicum, New York Institute of Technology. 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.

Hixson, M. D. (1999). The use of curriculum-based measurement to evaluate the effects of a remedial education program (Doctoral dissertation, Western Michigan University, 1999). *Dissertation Abstracts International*, *60*(06B), 53–2930. 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.

Holdsworth, P. (1984). *Corrective Reading* tested in UK. *ADI News, 4*(2), 1–4. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Additional source:

SRA/McGraw-Hill (n.d). *High I.Q. students*. Desoto, TX: Author.

Houchins, D. E., Jolivette, K., Krezmien, M. P., & Baltodano, H. M. (2008). A multi-state study examining the impact of explicit reading instruction with incarcerated students. *Journal of Correctional Education, 59*(1), 65–85. The study is ineligible for review because it does not use a comparison group.

Huang, L. V., Nelson, R. B., & Nelson, D. (2008). Increasing reading fluency through student-directed repeated reading and feedback. *California School Psychologist, 13*, 33–40. 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.

Humphries, T., Neufeld, M., Johnson, C., Engels, K., & McKay, R. (2005). A pilot study of the effect of Direct Instruction programming on the academic performance of students with intractable epilepsy. *Epilepsy & Behavior, 6*(3), 405–412. The study is ineligible for review because it does not use a comparison group.

Hurst, M., & Jolivette, K. (2006). Effects of private versus public assessment on the reading fluency of middle school students with mild disabilities. *Education and Training in Developmental Disabilities, 41*(2), 185. 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.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*(4), 437–447. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Kaiser, S., Palumbo, K., Bialozor, R. C., & McLaughlin, T. F. (1989). The effects of Direct Instruction with rural remedial students: A brief report. *Reading Improvement, 26*(1), 88–93. The study is ineligible for review because it does not use a comparison group.

Kalisek, A. M. (2004). The effects of a middle school *Corrective Reading* intervention on high school passage rate (Doctoral dissertation, University of La Verne, 2004). *Dissertation Abstracts International*, *65*(06A), 147–2138. 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.

Kaniuka, T. S. (1997). Impact of improving student academic achievement on teachers' instructional expectations and decision-making: Implications for the school reform process (Doctoral dissertation, East Carolina University, 1997). *Dissertation Abstracts International, 58*(02A), 255–352. 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.

Kasendorf, S., & McQuaid, P. (1987). *Corrective Reading* evaluation study. *Direct Instruction News, 9*. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Additional source:

SRA/McGraw-Hill. (n.d.). *Poor readers*. Desoto, TX: Author.

Keel, M. C., Frederick, L. D., Hughes, T. A., & Owens, S. H. (1999). Using paraprofessionals to deliver direct instruction reading programs. *Effective School Practices, 18*(2), 16–22. The study is ineligible for review because it does not use a comparison group.

Kelso, P. F. H. (2002). *The effectiveness of the SRA* Corrective Reading *curriculum in increasing reading level achievement of eighth grade students at Eagles' Landing Middle School*. Unpublished educational specialist's thesis, University of West Georgia, Carrollton. The study is ineligible for review because it does not use a comparison group.

References *(continued)*

Kinder, D., Kubina, R., & Marchand-Martella, N. E. (2005). Special education and Direct Instruction: An effective combination. *Journal of Direct Instruction, 5*(1), 1–36. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Kirby, V. (2007). Corrective Reading *intervention programme* 2006–2007. Kent, UK: New Brompton College. The study is ineligible for review because it does not use a comparison group.

Knudsen, S. L. (2006). Corrective Reading *program evaluation for struggling readers in third grade*. Unpublished master's thesis, Eastern Washington University, Cheney. The study is ineligible for review because it does not use a comparison group.

Kostewicz, D. E., & Kubina, R. M. (2008). The National Reading Panel guidepost: A review of reading outcome measures for students with emotional and behavioral disorders. *Behavioral Disorders, 33*(2), 62–74. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Krezmien, M. P., & Mulcahy, C. A. (2008). Literacy and delinquency: Current status of reading interventions with detained and incarcerated youth. *Reading & Writing Quarterly, 24*(2), 219–238. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Legault, A., Maloney, M., & Giroux, N. (2001). Learning rates with Direct Instruction, Precision Teaching and the *Corrective Reading* series. *Journal of Precision Teaching & Celeration, 17*(2), 89–91. The study is ineligible for review because it does not use a comparison group.

Lewis, A. (1982). An experimental evaluation of a Direct Instruction programme (*Corrective Reading*) with remedial readers in a comprehensive school. *Educational Psychology, 2*(2), 121–135. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Lenz, B. K., Ehren, B., & Deshler, D. D. (2005). *A framework for improving adolescent literacy*. Lawrence, KS: University of Kansas Center for Research on Learning. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Ligas, M. R. (2002). Evaluation of Broward County Alliance of Quality Schools project. *Journal of Education for Students Placed At Risk, 7*(2), 117–139. The study does not meet WWC evidence standards because the measures of effectiveness cannot be attributed solely to the intervention—the intervention was combined with another intervention.

Lingo, A. S. (2003). Effects of *Corrective Reading* on the reading abilities and classroom behaviors of middle school students with reading deficits and challenging behavior (Doctoral dissertation, University of Kentucky, 2003). *Dissertation Abstracts International, 64*(07A), 105–2446. 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.

Additional source:

Lingo, A. S., Slaton, D. B., & Jolivette, K. (2006). Effects of *Corrective Reading* on the reading abilities and classroom behaviors of middle school students with reading deficits and challenging behavior. *Behavioral Disorders, 31*(3), 265–283.

Lloyd, J., Epstein, M. H., & Cullinan, D. (1981). Direct teaching for learning disabilities. In J. Gottlieb & S. S. Strichart (Eds.), *Developmental theory and research in learning disabilities* (pp. 278–309). Baltimore: University Park Press. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Lovett, M. W., De Palma, M., Frijters, J., Steinbach, K., Temple, M., Benson, N., et al. (2008). Interventions for reading difficulties. *Journal of Learning Disabilities, 41*(4), 333–352. 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.

References *(continued)*

Malmgren, K. W., & Leone, P. E. (2000). Effects of a short-term auxiliary reading program on the reading skills of incarcerated youth. *Education & Treatment of Children, 23*, 239–247. The study is ineligible for review because it does not use a comparison group.

Marchand-Martella, N. E., & Martella, R. C. (2002). An overview and research summary of peer-delivered *Corrective Reading* instruction. *Behavior Analysis Today, 3*, 213–220. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Marchand-Martella, N., Martella, R. C., Bettis, D. F., & Blakely, M. R. (2004). Project PALS: A description of a high school–based tutorial program using *Corrective Reading* and peer-delivered instruction. *Reading & Writing Quarterly, 20*(2), 179–201. The study is ineligible for review because it does not use a comparison group.

Marchand-Martella, N., Martella, R., & Przychodzin-Havis, A. (n.d.). T*he research base and validation of SRA's* Corrective Reading *program*. Desoto, TX: SRA/McGraw-Hill. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Martin, C. (2006). *The effects of* Corrective Reading Decoding *combined with Spelling Through Morphographs on reading comprehension achievement of seventh-grade students*. Unpublished master's thesis, California State University–Stanislaus. The study is ineligible for review because it does not use a comparison group.

McMahon, K. S. (2002). *The effectiveness of the* Corrective Reading *program with middle school remedial readers*. Unpublished educational specialist's thesis, University of West Georgia, Carrollton. The study is ineligible for review because it does not use a comparison group.

Mitchell, S., & Wile, N. (2001). *2001 literacy program evaluation: A report of the evaluation of literacy programs in elementary and middle schools*. Portland, OR: Portland Public Schools, Research and Evaluation Department. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Nelson, J., & Herber, H. L. (1981). *A positive approach to assessment and correction of reading difficulties in middle and secondary schools*. Washington, DC: U.S. Department of Education. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

New York City Board of Education, Office of Educational Assessment. (1987). *Chapter 1* Corrective Reading *program 1985–1986: OEA evaluation report*. Brooklyn, NY: Author. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

New York City Board of Education, Office of Research, Evaluation and Assessment. (1990). *Chapter 1 clinical and guidance program 1988–89: Evaluation section report*. Brooklyn, NY: Author. The study is ineligible for review because it does not use a comparison group.

New York City Board of Education, *Office of Research, Evaluation and Assessment. (1991). Chapter 1 nonpublic school* Corrective Reading *program evaluation report 1989–90*: *OREA report*. Brooklyn, NY: Author. The study is ineligible for review because it does not use a comparison group.

Niemi, M. J. (2005). *Effects of* Corrective Reading Decoding *homework on fluency*. Unpublished master's thesis, University of Kentucky, Lexington. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Novak, N. L. (1997). *Reading achievement comparison of special education students instructed through the SRA* Corrective Reading *program*. Unpublished doctoral dissertation, State University of New York, Brockport. The study is ineligible for review because it does not use a comparison group.

Pfender, R. A. (2007). *The effects of a* Corrective Reading *program in a middle school*. Unpublished master's thesis, Gratz College, Melrose Park, PA. The study is ineligible for review because it does not use a comparison group.

Polloway, E. A. (1986). *Corrective Reading* program: An analysis of effectiveness with learning disabled and mentally retarded

students. *RASE: Remedial & Special Education, 7*(4), 41–47. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Przychodzin-Havis, A. M., Marchand-Martella, N. E., Martella, R. C., Miller, D. A., Warner, L., Leonard, B., et al. (2005). An analysis of *Corrective Reading* research. *Journal of Direct Instruction, 5*(1), 37. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Ramsch, M. (1988). *A comparison of the secondary* Corrective Reading *sequence approach, and the basal textbook approach as applied to a learning disabled child in the Petersburg public school system*. Unpublished master's thesis, Virginia State University, Petersburg. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Reed, D. K. (2008). A synthesis of morphology interventions and effects on reading outcomes for students in grades K–12. *Learning Disabilities Research & Practice, 23*(1), 36–49. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Regional Educational Laboratory Program. (2008). *Technical brief: Status of large-scale assessment in the Pacific region* (REL 2008-No. 003). Washington, DC: Author. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Ridout, J. L. (2002). *Effects of the SRA* Corrective Reading *program on the reading performance of eighth grade students with learning disabilities*. Unpublished master's thesis, Virginia State University, Petersburg. The study is ineligible for review because it does not use a comparison group.

Ross, D. (1998). *Competing theories: Teach comprehension or teach decoding to improve reading comprehension of older poor readers*. Unpublished manuscript. 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:

SRA/McGraw-Hill. (n.d.). *Remedial readers, Canada*. Desoto, TX: Author.

Rozalski, M. E. (2009). *Response to intervention: A rural high school's attempt to improve reading achievement*. Retrieved March 29, 2009, from *CEC Today*: http://www.cec.sped.org. 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.

Ruchti, K. R. (2005). Direct instruction of decoding skills: Effects on fluency rate of two learning disabled students (Master's thesis, Southwest Minnesota State University, 2005). *Masters Abstracts International, 44*(04), 44–1590. The study is ineligible for review because it does not use a comparison group.

Samuels, S. J., & Pearson, P. D. (Eds.). (1988). *Changing school reading programs: Principles and case studies*. Newark, DE: International Reading Association. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Scarlato, M. C., & Asahara, E. (2004). Effects of *Corrective Reading* in a residential treatment facility for adjudicated youth. *Journal of Direct Instruction, 4*(2), 211. 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.

Shippen, M. E. (2008). A pilot study of the efficacy of two adult basic literacy programs for incarcerated males. *Journal of Correctional Education, 59*(4), 339–347. 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.

Shippen, M., Houchins, D., Steventon, C., & Sartor, D. (2005). A comparison of two direct instruction reading programs for urban middle school students. *Remedial and Special Education, 26*(3), 175–182. The study does not meet WWC evidence standards because it is a randomized controlled trial in which

the combination of overall and differential attrition rates exceeds WWC standards for this area, and the subsequent analytic intervention and comparison groups are not shown to be equivalent.

Short, C., Marchand-Martella, N. E., Martella, R. C., & Ebey, T. L. (1999). The benefits of being high school *Corrective Reading* peer instructors. *Effective School Practices, 18*(2), 23–29. The study is ineligible for review because it does not use a comparison group.

Slatton, D. (2006). Effects of *Corrective Reading* on the reading abilities and classroom behaviors of middle school students with reading deficits and challenging behavior. *Behavioral Disorders, 31*(3), 265–283. 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.

Slavin, R. E., & Cheung, A. (*2003). Effective reading programs for English language learners. A best-evidence synthesis*. Report No. 66. Baltimore, MD: Center for Research on the Education of Students Placed at Risk (CRESPAR). 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.

Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly, 43*(3), 290–322. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Soehnlein, J. (2006). *Effects of tutor-implemented* Corrective Reading *and Reading Success on the reading achievement of high school students.* Unpublished master's thesis, Eastern Washington University, Cheney. The study is ineligible for review because it does not use a comparison group.

Somerville, D. E., & Leach, D. J. (1988). Direct or indirect instruction? An evaluation of three types of intervention programme for assisting students with specific reading difficulties. *Educational Research, 30*(1), 46–53. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Sommers, J. (1995). Seven-year overview of direct instruction programs used in basic skills classes at Big Piney Middle School. *Effective School Practices, 14*(4), 29–32. The study is ineligible for review because it does not use a comparison group.

SRA/McGraw-Hill. (2006). *Reading Mastery,* Corrective Reading *help students with disabilities achieve significant academic growth*. Desoto, TX: Author. 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.

SRA/McGraw-Hill. (n.d.). *At risk students (Lee County, Alabama)*. Desoto, TX: Author. The study is ineligible for review because it does not use a comparison group.

SRA/McGraw-Hill. (n.d.). *Results with* Corrective Reading *Direct Instruction in middle school and high school*. Desoto, TX: Author. The study is ineligible for review because it does not use a comparison group.

Stephens, M. A. (1993). *Developing and implementing a curriculum and instructional program to improve reading achievement of middle-grade students with learning disabilities in a rural school district*. Unpublished doctoral research project, Nova University, Ft. Lauderdale, FL. The study is ineligible for review because it does not use a comparison group.

Steventon, C. E., & Fredrick, L. D. (2003). The effects of repeated readings on student performance in the *Corrective Reading* program. *Journal of Direct Instruction, 3*(1), 17. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Strong, A. C., Wehby, J. H., Falk, K. B., & Lane, K. L. (2004). The impact of a structured reading curriculum and repeated reading on the performance of junior high students with emotional and behavioral disorders. *School Psychology Review, 33*(4), 561–581. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Sulgrove, M. K., & McLaughlin, T. F. (2004). The effects of an additional timed reading on reading rate. *Journal of Precision*

Teaching & Celeration, 20(1), 9–16. The study is ineligible for review because it does not use a comparison group.

Syverud, S. M. (2004). Lingering questions regarding the transfer effects of improvements in oral reading fluency (Doctoral dissertation, University of Wisconsin–Madison, 2004). *Dissertation Abstracts International*, *65*(08A), 104–2952. 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.

Thorne, M. (1978). Payment for reading: The use of the *Corrective Reading* scheme with junior maladjusted boys. *Remedial Education, 13*(2), 87–90. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Additional source:

SRA/McGraw-Hill. (n.d.). *Individuals with severe behavioral problems*. Desoto, TX: Author.

Uhry, J. K., & Clark, D. B. (2004). *Dyslexia: Theory & practice of instruction* (3rd ed.). Austin, TX: PRO-ED. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Vitale, M., & Joseph, B. (2008). Broadening the institutional value of direct instruction implemented in a low-SES elementary school: Implications for scale-up and school reform. *Journal of Direct Instruction, 8*(1), 1–18. The study is ineligible for review because it does not use a comparison group.

Vitale, M. R., Medland, M. B., Romance, N., & Weaver, H. P. (1993). Accelerating reading and thinking skills of low-achieving elementary students: Implications for curricular change. *Effective School Practices, 12*(1), 26–31. 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.

Watson, T., & Hempenstall, K. (2008). Effects of a computer based beginning reading program on young children. *Australasian Journal of Educational Technology, 24*(3), 258–274. 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.

Wedman, J. M., & Robinson, R. D. (1989). A survey of diagnostic/*Corrective Reading* instruction and practices in relation to the interactive reading process. *Reading Horizons, 29*(3), 167. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Werner, D. H. (2005). A study to determine the relationship of the Direct Instruction program *Corrective Reading* on Terra Nova tests scores in one school system in East Tennessee. *Dissertation Abstracts International, 69*(3-A), 881. 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.

WestEd. (2004). *Innovations in education: Successful charter schools*. Washington, DC: U.S. Department of Education. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Westphal, R. C., Jr., & Tuss, P. (2002). *Evaluation of the 2001–02* Corrective Reading *program*. San Juan, CA: Department of Accountability and Organizational Evaluation. The study is ineligible for review because it does not use a comparison group.

Wilson, B. (2000). Educators' views of implementing Direct Instruction curricula: Connections to students with disabilities (Doctoral dissertation, West Virginia University, 2000). *Dissertation Abstracts International, 62*(04A), 239–1318. 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.

Worner, L. J. (1989). Corrective Reading: *An effective method for teaching severely learning disabled elementary students*. Unpublished master's thesis, Moorhead State University, MN. 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.

References *(continued)*

Yawn, C. D. (2008). *Effects of peer-mediated Direct Instruction and repeated reading on the reading skills of incarcerated juveniles with disabilities*. Unpublished doctoral dissertation, The Ohio State University, Columbus. The study is ineligible for review because it does not use a comparison group.

Yevoli, C. (1993). *Corrective strategies in reading for at-risk community college students*. Unpublished master's thesis, Long Island University, Brookville, NY. 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.

Appendix

Appendix A1 Study characteristics: Torgesen et al. (2006)

Characteristic	Description
Study citation	Torgesen, J., Myers, D., Schirm, A., Stuart, E., Vartivarian, S., Mansfield, W., et al. (2006). *National assessment of Title I. Interim report. Volume II: Closing the reading gap: First year findings from a randomized trial of four reading interventions for striving readers*. Washington, DC: National Center for Education Evaluation and Regional Assistance. ***Additional source:*** Torgesen, J., Schirm, A., Castner, L., Vartivarian, S., Mansfield, W., Myers, D., et al. (2007). *National assessment of Title I. Final report. Volume II: Closing the reading gap: Findings from a randomized trial of four reading interventions for striving readers* (NCEE 2008-4013). Washington, DC: National Center for Education Evaluation and Regional Assistance.
Participants	The study design was based on random assignment of 32 school units,[1] formed from a pool of 52 schools, to one of four interventions (*Corrective Reading, Kaplan SpellRead, Failure Free Reading*, and *Wilson Reading*). Within each school, students were randomly assigned to the treatment group that would receive the intervention assigned to its school or to the control group that would receive the standard reading curriculum. This report focuses on schools assigned to *Corrective Reading* and on findings for 5th graders (as specified by the Adolescent Literacy review protocol). At the time of the analysis, the sample relevant to this review included 86 fifth-grade students (55 in *Corrective Reading* and 31 in the control group) in seven school units. The number of 5th-grade students at baseline was not reported.[2] Students were eligible for participation if their teacher identified them as a struggling reader and if they scored at or below the 30th percentile on a word-level reading test and at or above the 5th percentile on a vocabulary test. On average, at baseline, students scored about one-half to one standard deviation below national norms on measures used to assess their ability to decode words. About 51% of the intervention group students were females, compared to 36% in the control group. About 41% of the intervention group students were eligible for free or reduced-price lunch programs, equal to 41% of the students in the control group.
Setting	The analysis sample included seven school units in the Allegheny Intermediate Unit (AIU), outside Pittsburgh, Pennsylvania. The AIU consisted of 42 school districts.
Intervention	The decoding component of *Corrective Reading* was implemented by nine teachers beginning in the first week in November 2003 through the first week in May 2004. The comprehension component was not implemented. The intervention was administered to students in groups of three that were heterogeneous with regard to students' basic reading skills. The average skills of the students in each of the instructional groups determined the pace of instruction. Implementation fidelity was determined by reading program trainers who observed the teachers and coached them over a period of months, project coordinators who observed a sample of instructional sessions, and ratings based on a sample of videotaped sessions. Implementation was rated as acceptable. The decoding component used in the study included four levels—A, B1, B2, and C. Placement testing was used to start each group at the appropriate level. The lessons provided during the study clustered in levels B1 and B2. For those groups that progressed to level C, explicit vocabulary instruction was not provided. Over a six-month period, students received a total of about 90 hours of instruction. Students received *Corrective Reading* instruction five days a week in sessions that were approximately 55 minutes long. The study reported student outcomes after six months of program implementation. Additional findings reflecting students' outcomes one year after the end of the implementation of the intervention can be found in Appendices A4.1–A4.3.
Comparison	The control group students received their regular reading instruction, which included typical classroom instruction and, in many cases, other services (such as another pull-out program). Across four interventions, the control group students had fewer small-group instructional hours and average weekly hours of total reading instruction than the intervention group students.

(continued)

Appendix A1 Study characteristics: Torgesen et al. (2006) *(continued)*

Characteristic	Description
Primary outcomes and measurement	The primary outcome measures in the alphabetics domain were the Word Identification and Word Attack subtests of the Woodcock Reading Mastery Test–Revised (WRMT-R) and the Phonetic Decoding Efficiency and Sight Word Efficiency subtests of the Test of Word Reading Efficiency (TOWRE). The primary measure in the reading fluency domain was the Oral Reading Fluency test (also referred to as AIMSweb). The primary measures in the comprehension domain were the WRMT-R: Comprehension subtest and the Group Reading Assessment and Diagnostic Evaluation (GRADE): Passage Comprehension subtest. For a more detailed description of these outcome measures, see Appendices A2.1–A2.3.
Staff/teacher training	Professional development on how to use *Corrective Reading* included training and coaching by *Corrective Reading* program staff, teachers' independent study of program materials, and telephone conferences between teachers and *Corrective Reading* staff. On average, throughout the course of the study, the *Corrective Reading* intervention group teachers participated in 70.8 professional development hours specifically related to using *Corrective Reading* (32.8 hours were initial training in use of the program, 26.4 hours were spent in a practice phase, and 11.6 hours occurred during the six-month period in which teachers were using *Corrective Reading*).

1. A school unit consists of several schools partnering so that the cluster included two 3rd-grade and two 5th-grade instructional groups.
2. The study reported that 10 students in the intervention group and no students in the control group were lost to analysis. However, it is not clear if those students were in 5th grade or were part of the sample of 3rd-grade students that was also examined in this study. The 3rd-grade sample that was included in this study is not reviewed in this report because it is outside the scope of the review.

Appendix A2.1 Outcome measures for the alphabetics domain

Outcome measure	Description
Phonics construct	
Woodcock Reading Mastery Test–Revised (WRMT-R): Word Attack subtest	This standardized test measures phonemic decoding skills by asking students to pronounce printed pseudo-words. Students are aware that the words are not real (as cited in Torgesen et al., 2006).
Woodcock Reading Mastery Test–Revised (WRMT-R): Word Identification subtest	The word identification subtest is a test of decoding skills. The standardized test requires children to pronounce real words from a list of increasing difficulty (as cited in Torgesen et al., 2006).
Test of Word Reading Efficiency (TOWRE): Phonemic Decoding Efficiency (PDE) subtest	The TOWRE is a standardized, nationally normed measure. The PDE subtest measures the number of nonwords of increasing difficulty that students can pronounce within 45 seconds (as cited in Torgesen et al., 2006).
Test of Word Reading Efficiency (TOWRE): Sight Word Efficiency (SWE) subtest	The TOWRE is a standardized, nationally normed measure. The SWE subtest measures the number of real words of increasing difficulty that students can pronounce within 45 seconds (as cited in Torgesen et al., 2006).

Appendix A2.2 Outcome measures for the reading fluency domain

Outcome measure	Description
Oral Reading Fluency assessment	This test (also referred to as AIMSweb) measures the number of words correct per minute (WCPM) that students read using three brief grade-level-appropriate passages. These passages contain both fiction and nonfiction text. The norms for this test are updated by Edformation each school year (as cited in Torgesen et al., 2006).

Appendix A2.3 Outcome measures for the comprehension domain

Outcome measure	Description
Reading comprehension construct	
Woodcock Reading Mastery Test–Revised (WRMT-R): Passage Comprehension subtest	In this standardized test, comprehension is measured by having students read silently and fill in missing words in a short paragraph (as cited in Torgesen et al., 2006).
Group Reading Assessment and Diagnostic Evaluation (GRADE): Passage Comprehension subtest	The GRADE is a norm-referenced reading assessment that can be used with students at any level. The GRADE has four subtests: (1) Vocabulary, (2) Sentence Comprehension, (3) Passage Comprehension, and (4) Listening Comprehension. The Passage Comprehension subtest includes a passage of text and corresponding multiple-choice comprehension questions (as cited in Torgesen et al., 2006).

Appendix A3.1 Summary of study findings included in the rating for the alphabetics domain[1]

Outcome measure	Study sample	Sample size (school units/ students)	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: *Corrective Reading* group	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: Comparison group	WWC calculations: Mean difference[4] (*Corrective Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Torgesen et al., 2006[8]								
WRMT-R: Word Attack subtest	Grade 5	7/86	97.40 (15.00)	95.50 (15.00)	1.90	0.13	ns	+5
WRMT-R: Word Identification subtest	Grade 5	7/86	92.90 (15.00)	92.60 (15.00)	0.30	0.02	ns	+1
TOWRE: Phonemic Decoding Efficiency (PDE) subtest	Grade 5	7/86	87.30 (15.00)	85.40 (15.00)	1.90	0.13	ns	+5
TOWRE: Sight Word Efficiency (SWE) subtest	Grade 5	7/86	88.70 (15.00)	86.50 (15.00)	2.20	0.15	ns	+6
Domain average for alphabetics (Torgesen et al., 2006)[9]						**0.10**	**ns**	**+4**

ns = not statistically significant
WRMT-R = Woodcock Reading Mastery Test–Revised
TOWRE = Test of Word Reading Efficiency

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the alphabetics domain. Follow-up findings from the same studies are not included in these ratings but are reported in Appendix A4.1. Torgesen et al. (2006) also included subgroup analyses by initial skill level (WRMT-R Word Attack subtest and Peabody Picture Vocabulary Test [PPVT]) and socioeconomic status. The study reported that *Corrective Reading* had statistically significant positive effects on TOWRE Sight Word Efficiency for students with low initial Word Attack scores and low initial PPVT scores, and for students who were eligible for free or reduced-price lunch programs. The study also found statistically significant positive effects on WRMT-R Word Attack scores for students with low initial PPVT scores. No other differences were found between subgroups of students for outcomes in the alphabetics domain.
2. For Torgesen et al. (2006), the mean outcomes were computed using information reported in the paper. For the control group, the mean outcome is the control group baseline mean standard score plus the control group gain. For the intervention group, the mean outcome is the control group baseline mean standard score plus the control group gain plus the impact of the intervention.
3. 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. The standard deviations in the Torgesen et al. (2006) study were the population standard deviations for these standardized outcomes.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the control group.
5. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
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 and that of the average student in the control condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results for the intervention group.
8. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Torgesen et al. (2006) and the alphabetics domain, no corrections for clustering were needed because the authors adjusted for clustering, and no correction for multiple comparisons was needed because there were no statistically significant findings in this domain.
9. This row provides the study average, which in this instance is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A3.2 Summary of study findings included in the rating for the reading fluency domain[1]

Outcome measure	Study sample	Sample size (school units/ students)	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: *Corrective Reading* group	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: Comparison group	WWC calculations: Mean difference[4] (*Corrective Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Torgesen et al., 2006[8]								
Oral Reading Fluency test	Grade 5	7/86	96.80 (47.00)	91.90 (47.00)	4.90	0.10	ns	+4
Domain average for reading fluency (Torgesen et al., 2006)[9]						**0.10**	**ns**	**+4**

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the reading fluency domain. Follow-up findings from the same studies are not included in these ratings but are reported in Appendix A4.2. The study also included subgroup analyses by initial skill level (WRMT-R Word Attack subtest and Peabody Picture Vocabulary Test [PPVT]) and socioeconomic status. No differences were found between subgroups of students for the reading fluency outcome.
2. For Torgesen et al. (2006), the mean outcomes were computed using information reported in the paper. For the control group, the mean outcome is the control group baseline mean standard score plus the control group gain. For the intervention group, the mean outcome is the control group baseline mean standard score plus the control group gain plus the impact of the intervention.
3. 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. The standard deviations in the Torgesen et al. (2006) study were the population standard deviations for these standardized outcomes.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the control group.
5. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
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 and that of the average student in the control condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results for the intervention group.
8. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Torgesen et al. (2006) and the reading fluency domain, no corrections for clustering were needed because the authors adjusted for clustering, and no correction for multiple comparisons was needed because there is only one outcome in this domain.
9. This row provides the study average, which in this instance is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A3.3 Summary of study findings included in the rating for the comprehension domain[1]

Outcome measure	Study sample	Sample size (school units/ students)	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: *Corrective Reading* group	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: Comparison group	WWC calculations: Mean difference[4] (*Corrective Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Torgesen et al., 2006[8]								
WRMT-R: Passage Comprehension subtest	Grade 5	7/86	93.80 (15.00)	92.00 (15.00)	1.80	0.12	ns	+5
GRADE: Passage Comprehension subtest	Grade 5	7/86	96.30 (15.00)	96.00 (15.00)	0.30	0.02	ns	+1
Domain average for comprehension (Torgesen et al., 2006)[9]						**0.07**	**ns**	**+3**

ns = not statistically significant
WRMT-R = Woodcock Reading Mastery Test–Revised
GRADE = Group Reading Assessment and Diagnostic Evaluation

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the comprehension domain. Follow-up findings from the same studies are not included in these ratings but are reported in Appendix A4.3. The study also included subgroup analyses by initial skill level (WRMT-R Word Attack subtest and Peabody Picture Vocabulary Test [PPVT]) and socioeconomic status. No differences were found between subgroups of students for outcomes in the comprehension domain.
2. For Torgesen et al. (2006), the mean outcomes were computed using information reported in the paper. For the control group, the mean outcome is the control group baseline mean standard score plus the control group gain. For the intervention group, the mean outcome is the control group baseline mean standard score plus the control group gain plus the impact of the intervention.
3. 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. The standard deviations in the Torgesen et al. (2006) study were the population standard deviations for these standardized outcomes.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the control group.
5. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
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 and that of the average student in the control condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results for the intervention group.
8. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Torgesen et al. (2006) and the comprehension domain, no corrections for clustering were needed because the authors adjusted for clustering. No correction for multiple comparisons was needed because there were no statistically significant findings in this domain.
9. This row provides the study average, which in this instance is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A4.1 Summary of follow-up findings for the alphabetics domain[1]

Outcome measure	Study sample	Sample size (school units/ students)	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: *Corrective Reading* group	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: Comparison group	WWC calculations: Mean difference[4] (*Corrective Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Torgensen, et al., 2007[8]								
WRMT-R: Word Attack subtest	Grade 5	7/84	98.60 (15.00)	98.40 (15.00)	0.20	0.01	ns	+1
WRMT-R: Word Identification subtest	Grade 5	7/84	92.10 (15.00)	94.00 (15.00)	–1.90	–0.13	ns	–5
TOWRE: Phonemic Decoding Efficiency (PDE) subtest	Grade 5	7/84	90.00 (15.00)	88.60 (15.00)	1.40	0.09	ns	+4
TOWRE: Sight Word Efficiency (SWE) subtest	Grade 5	7/84	87.10 (15.00)	87.50 (15.00)	–0.40	–0.03	ns	–1

ns = not statistically significant
WRMT-R = Woodcock Reading Mastery Test–Revised
TOWRE = Test of Word Reading Efficiency

1. This appendix presents findings from data collected one year after the end of the implementation of the intervention for measures that fall in the alphabetics domain. Data that reflected students' exposure to six months of the intervention were used for rating purposes and are presented in Appendix A3.1. Torgesen et al. (2007) also included subgroup analyses by initial skill level (WRMT-R Word Attack subtest and Peabody Picture Vocabulary Test [PPVT]) and socioeconomic status. The study reported that *Corrective Reading* had a statistically significant positive effect on the TOWRE PDE and TOWRE SWE for students eligible for free or reduced-price lunch programs. The study also reported that *Corrective Reading* had a statistically significant positive effect on the TOWRE SWE for students with low initial PPVT scores and on the TOWRE PDE for students with low initial Word Attack scores but high initial PPVT scores. Additional subgroup analyses found statistically significant positive effects on Word Attack for students with low initial Word Attack scores and for students with low scores on both baseline assessments (Word Attack and PPVT). No other differences were reported between subgroups of students for outcomes in the alphabetics domain.
2. For Torgesen et al. (2007), the mean outcomes were computed using information reported in the paper. For the control group, the mean outcome is the control group baseline mean standard score plus the control group gain. For the intervention group, the mean outcome is the control group baseline mean standard score plus the control group gain plus the impact of the intervention.
3. 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. The standard deviations in the Torgesen et al. (2007) study were the population standard deviations for these standardized outcomes.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the control group.
5. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
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 and that of the average student in the control 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, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Torgesen et al. (2007) and the alphabetics domain, no corrections for clustering were needed because the authors adjusted for clustering, and no correction for multiple comparisons was needed because there were no statistically significant findings in this domain.

Appendix A4.2 Summary of follow-up findings for the reading fluency domain[1]

Outcome measure	Study sample	Sample size (school units/ students)	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: *Corrective Reading* group	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: Comparison group	WWC calculations: Mean difference[4] (*Corrective Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Torgesen et al., 2007[8]								
Oral Reading Fluency test	Grade 5	7/84	102.10 (47.00)	107.40 (47.00)	–5.30	–0.11	ns	–4

ns = not statistically significant

1. This appendix presents findings from data collected one year after the end of the implementation of the intervention for measures that fall in the reading fluency domain. Data that reflected students' exposure to six months of the intervention were used for rating purposes and are presented in Appendix A3.2. Torgesen et al. (2007) also included subgroup analyses by initial skill level (WRMT-R Word Attack subtest and Peabody Picture Vocabulary Test [PPVT]) and socioeconomic status. The study reported that *Corrective Reading* had a statistically significant negative effect on the Oral Reading Fluency test for students with high initial PPVT scores, for students who had both high initial word attack scores and high initial PPVT scores, and for students who were not eligible for free or reduced-price lunch programs. No other differences were found between subgroups of students for the reading fluency outcome.
2. For Torgesen et al. (2007), the mean outcomes were computed using information reported in the paper. For the control group, the mean outcome is the control group baseline mean standard score plus the control group gain. For the intervention group, the mean outcome is the control group baseline mean standard score plus the control group gain plus the impact of the intervention.
3. 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. The standard deviations in the Torgesen et al. (2007) study were the population standard deviations for these standardized outcomes.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the control group.
5. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
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 and that of the average student in the control 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, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Torgesen et al. (2007) and the reading fluency domain, no corrections for clustering were needed because the authors adjusted for clustering, and no correction for multiple comparisons was needed because there is only one outcome in this domain.

Appendix A4.3 Summary of follow-up findings for the comprehension domain[1]

Outcome measure	Study sample	Sample size (school units/ students)	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: *Corrective Reading* group	Authors' findings from the study: Mean outcome[2] (standard deviation)[3]: Comparison group	WWC calculations: Mean difference[4] (*Corrective Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Torgesen et al., 2007[8]								
WRMT-R: Passage Comprehension	Grade 5	7/84	93.50 (15.00)	95.60 (15.00)	–2.10	–0.14	ns	–6
GRADE: Passage Comprehension	Grade 5	7/84	91.60 (15.00)	91.60 (15.00)	0	0	ns	0

ns = not statistically significant

WRMT-R = Woodcock Reading Mastery Test–Revised

GRADE = Group Reading Assessment and Diagnostic Evaluation

1. This appendix presents findings from data collected one year after the end of the implementation of the intervention for measures that fall in the comprehension domain. Data that reflected students' exposure to six months of the intervention were used for rating purposes and are presented in Appendix A3.3. Torgesen et al. (2007) also included subgroup analyses by initial skill level (WRMT-R Word Attack subtest and Peabody Picture Vocabulary Test [PPVT]) and socioeconomic status. The study reported that *Corrective Reading* had a statistically significant negative effect on both comprehension outcomes for students with high initial level of Word Attack scores. The difference was also negative and statistically significant on the GRADE outcome for students with high initial level of Word Attack and PPVT scores, but positive and statistically significant for students with low initial PPVT scores. The study also found a statistically significant positive effect on the WRMT Passage Comprehension for students with low skill level on both baseline assessments: Word Attack and PPVT. No other differences were reported between subgroups of students for outcomes in the comprehension domain.
2. For Torgesen et al. (2007), the mean outcomes were computed using information reported in the paper. For the control group, the mean outcome is the control group baseline mean standard score plus the control group gain. For the intervention group, the mean outcome is the control group baseline mean standard score plus the control group gain plus the impact of the intervention.
3. 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. The standard deviations in the Torgesen et al. (2007) study were the population standard deviations for these standardized outcomes.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the control group.
5. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
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 and that of the average student in the control 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, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Torgesen et al. (2007) and the comprehension domain, no corrections for clustering were needed because the authors adjusted for clustering. No correction for multiple comparisons was needed because there were no statistically significant findings in this domain.

Appendix A5.1 *Corrective Reading* rating for the alphabetics domain

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

For the outcome domain of alphabetics, the WWC rated *Corrective Reading* as having no discernible effects for adolescent learners.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: No studies showing a statistically significant or substantively important effect, either *positive* or *negative*.

  Met. No studies 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 studies showed a statistically significant positive effect.

AND

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

  Met. No studies showed a statistically significant or substantively important negative effect.

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 studies 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 studies showed a statistically significant or substantively important negative effect. One study 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 studies showed a statistically significant or substantively important positive effect, and no studies showed a statistically significant or substantively important negative effect.

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 studies showed a statistically significant or substantively important effect. One study showed indeterminate effects.

(continued)

Appendix A5.1 *Corrective Reading* rating for the alphabetics domain *(continued)*

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

- Criterion 1: One study showing a statistically significant or substantively important *negative* effect and no studies showing a statistically significant or substantively important *positive* effect.

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

OR

- Criterion 2: Two or more studies showing statistically significant or substantively important *negative* effects, at least one study showing a statistically significant or substantively important *positive* effect, and more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

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

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 studies showed a statistically significant negative effect.

AND

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

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

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. For a complete description, see the WWC Procedures and Standards Handbook, Appendix E.

Appendix A5.2 *Corrective Reading* rating for the reading fluency domain

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

For the outcome domain of reading fluency, the WWC rated *Corrective Reading* as having no discernible effects for adolescent learners.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: No studies showing a statistically significant or substantively important effect, either *positive* or *negative*.

  Met. No studies 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 studies showed a statistically significant positive effect.

AND

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

  Met. No studies showed a statistically significant or substantively important negative effect.

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 studies 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 studies showed a statistically significant or substantively important negative effect. One study 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 studies showed a statistically significant or substantively important positive effect, and no studies showed a statistically significant or substantively important negative effect.

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 studies showed a statistically significant or substantively important effect. One study showed indeterminate effects.

(continued)

Appendix A5.2 *Corrective Reading* rating for the reading fluency domain *(continued)*

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

- Criterion 1: One study showing a statistically significant or substantively important *negative* effect and no studies showing a statistically significant or substantively important *positive* effect.

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

OR

- Criterion 2: Two or more studies showing statistically significant or substantively important *negative* effects, at least one study showing a statistically significant or substantively important *positive* effect, and more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

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

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 studies showed a statistically significant negative effect.

AND

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

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

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. For a complete description, see the WWC Procedures and Standards Handbook, Appendix E.

Appendix A5.3 *Corrective Reading* rating for the comprehension domain

The WWC rates an intervention's effects for 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 *Corrective Reading* as having no discernible effects for adolescent learners.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: No studies showing a statistically significant or substantively important effect, either *positive* or *negative*.

  Met. No studies 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 studies showed a statistically significant positive effect.

AND

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

  Met. No studies showed a statistically significant or substantively important negative effect.

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 studies 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 studies showed a statistically significant or substantively important negative effect. One study 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 studies showed a statistically significant or substantively important positive effect, and no studies showed a statistically significant or substantively important negative effect.

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 studies showed a statistically significant or substantively important effect. One study showed indeterminate effects.

(continued)

Appendix A5.3 *Corrective Reading* rating for the comprehension domain *(continued)*

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

- Criterion 1: One study showing a statistically significant or substantively important *negative* effect and no studies showing a statistically significant or substantively important *positive* effect.

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

OR

- Criterion 2: Two or more studies showing statistically significant or substantively important *negative* effects, at least one study showing a statistically significant or substantively important *positive* effect, and more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

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

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 studies showed a statistically significant negative effect.

AND

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

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

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. For a complete description, see the WWC Procedures and Standards Handbook, Appendix E.

Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence[1]
		Schools	Students	
Alphabetics	1	7 school units[2]	86	Small
Reading fluency	1	7 school units[2]	86	Small
Comprehension	1	7 school units[2]	86	Small
General literacy achievement	na	na	na	na

na = not applicable/not studied

1. A rating of "medium 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." For more details on the extent of evidence categorization, see the WWC Procedures and Standards Handbook, Appendix G.
2. A school unit consists of several schools and includes two 3rd-grade and two 5th-grade instructional groups. The exact number of schools participating in *Corrective Reading* is unknown.