What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Adolescent Literacy August 2010

Reading Mastery

Program Description[1]

Reading Mastery, one of several curriculum components that constitute the *Direct Instruction* curriculum from SRA/McGraw-Hill, is designed to provide systematic instruction in reading to students in grades K–6. *Reading Mastery*, which can be used as an intervention program for struggling readers, as a supplement to a school's core reading program, or as a stand-alone reading program, is available in three versions:

1. *Reading Mastery Classic* (for use in grades pre-K–3) aims to help beginning readers identify letter sounds, segment words into sounds, blend sounds into words, develop vocabulary, and begin to learn comprehension strategies.
2. *Reading Mastery Plus* (for grades K–6) has a language-arts focus with an emphasis on reading, writing, spelling, and language.
3. *Reading Mastery Signature Edition* (for use in grades K–5) includes three strands: (a) the Reading strand addresses phonemic awareness, phonics, word analysis, fluency, vocabulary, comprehension, spelling, decoding, and word recognition skills; (b) the Oral Language/Language Arts strand addresses oral language, communication, and writing skills; and (c) the Literature strand is designed to provide students with opportunities to read a variety of different types of text and to develop their vocabulary.

During the implementation of *Reading Mastery*, students are grouped with other students at a similar reading level, based on program placement tests. The program includes a continuous monitoring component.

Research[2]

One study of *Reading Mastery* that falls within the scope of the Adolescent Literacy review protocol meets What Works Clearinghouse (WWC) evidence standards, and one study meets WWC evidence standards with reservations. The studies included 361 students in grades 4 and 5, who attended schools in the midwestern and northwestern United States.[3]

1. The descriptive information for this program was obtained from a publicly available source: the program's website (http://www.mcgraw-hill.co.uk/sra/downloads/Reading%20Mastery/Reading%20Mastery%20Signature%20Edition%20Brochure.pdf, downloaded October 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 August 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.

Research[2] *(continued)*

Based on two studies, the WWC considers the extent of evidence for *Reading Mastery* on adolescent learners to be small for the reading fluency and comprehension domains. No studies that meet WWC evidence standards with or without reservations examined the effectiveness of *Reading Mastery* on adolescent learners in the alphabetics or general literacy achievement domains.

Effectiveness

Reading Mastery was found to have potentially positive effects on reading fluency and no discernible effects on comprehension for adolescent learners.

	Alphabetics	*Reading fluency*	*Comprehension*	*General literacy achievement*
Rating of effectiveness	na	Potentially positive effects	No discernible effects	na
Improvement index[4]	na	+19 percentile points	Average: –7 percentile points	na
	na	na	Range: –20 to +7 percentile points	na

na = not applicable

Additional program information

Developer and contact

Reading Mastery was originally called DISTAR (*Direct Instruction System for Teaching Arithmetic and Reading*). Early versions of *Reading Mastery* were developed during the 1960s and 1970s by Dr. Siegfried Engelmann as part of the *Direct Instruction* teaching model.[5] *Reading Mastery* is distributed by SRA/McGraw-Hill, 220 East Danieldale Road, DeSoto, TX 75115-2490. Web: www.sraonline.com. Email: SRA_CustomerService@mcgraw-hill.com. Telephone: (201) 512-0909.

Scope of use

Reading Mastery is appropriate for elementary-age children who are above, at, or below grade level in their reading performance. The program also can be used with English language learners and special education students.

Cost

Student materials include storybooks (grades pre-K–1) or textbooks (grades 2–6), workbooks, and test books. The cost per student ranges from $200 to $300 for the first year of implementation. A full set of teaching materials—a one-time purchase—costs between $650 and $1,000 for each grade level. Additional components include literature collections, Independent Readers, seatwork blackline masters, and Practice and Review CD-ROMs for student practice of skills taught in the program. SRA Teaching Tutor CD-ROMs supplement consultant-led professional development. Additional information on costs of training materials and workshops is available online.

4. These numbers show the average and range of student-level improvement indices for all findings across the study.
5. This program is also known as *Direct Instruction* using the *Reading Mastery* texts or *SRA Direct Instruction—Reading Mastery.*

Additional program information *(continued)*

Teaching

A typical 30- to 45-minute *Reading Mastery* lesson includes seven to nine short activities that encompass multiple strands of content, such as phonemic awareness, letter-sound correspondence, sounding out words, word recognition, vocabulary, oral reading fluency, or comprehension. The teaching routine repeated throughout the curriculum is composed of the following steps: modeling new content, providing guided practice, and implementing individual practice and application. Lesson scripts act as a guide for teachers. Signals and group responses are used to keep students involved and on task, and to control lesson pacing. The program typically spans one academic year.

Research

A total of 175 studies reviewed by the WWC investigated the effects of *Reading Mastery* on adolescent learners. One study (Stockard, 2010) is a randomized controlled trial that meets WWC evidence standards, and one study (Yu & Rachor, 2000) is a quasi-experimental design that meets WWC evidence standards with reservations. The remaining 173 studies do not meet either WWC evidence standards or eligibility screens.

Meets evidence standards

Stockard (2010) conducted a randomized controlled study that examined the effects of *Reading Mastery Signature Edition* on 4th graders in an elementary school in the midwestern United States. General education students were randomly assigned to a treatment or control condition using alternative assignment with random start technique (see Appendix A1.1 for more details on the random assignment method). Two pairs of teachers were randomly assigned to a treatment or control condition via a coin flip. The WWC based its effectiveness ratings on findings from comparisons of the 29 students in two classrooms who received *Reading Mastery Signature Edition* and the 28 control students in two classrooms who received the *Scott Foresman Basal Reading Program*. The study reported students' outcomes after five months of program implementation.

Meets evidence standards with reservations

Yu and Rachor (2000) conducted a retrospective quasi-experimental study that examined the effects of *Reading Mastery* on students from three grades in six schools in the northwestern United States (three schools implemented *Reading Mastery* and three did not). Each of the three *Reading Mastery* schools was matched (based on poverty level and percentage of minority students) with a comparison school that was not using *Reading Mastery*. Then, *Reading Mastery* students were matched to comparison-group students within each grade level on the basis of reading scores and demographic characteristics. The WWC based its effectiveness rating on findings from students in two grades.[6] The 4th-grade group consisted of 71 students who received *Reading Mastery* and 71 students in the comparison group who received standard reading instruction. The 5th-grade group consisted of 81 students in the *Reading Mastery* group and 81 students in the comparison group. The study reported students' outcomes after one and two years of program implementation.[7]

6. Findings for 3rd-grade students are outside the scope of the Adolescent Literacy review.
7. Two-year findings are considered for the effectiveness rating because these findings reflect the maximum exposure to the program. One-year findings are not included in this rating but are reported in Appendix A4.

Research *(continued)*

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.[8]

The WWC considers the extent of evidence for *Reading Mastery* to be small for the reading fluency and comprehension domains for adolescent learners. No studies that meet WWC evidence standards with or without reservations examined the effectiveness of *Reading Mastery* on adolescent learners in the alphabetics or general literacy achievement domains.

Effectiveness

Findings

The WWC review of Adolescent Literacy interventions addresses student outcomes in four domains: alphabetics, reading fluency, comprehension, and general literacy achievement. The studies included in this report cover two domains: 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 *Reading Mastery* on adolescent learners.[9]

Reading fluency. Stockard (2010) found statistically significant positive effects of *Reading Mastery* on the AIMS Web Curriculum-Based Measurement Words Read Correct for 4th graders. According to WWC calculations, the effect was not statistically significant, but it was large enough to be considered substantively important (i.e., an effect size of at least 0.25). Thus, for the reading fluency domain, one study showed substantively important positive effects.

Comprehension. Yu and Rachor (2000) found a statistically significant positive effect of *Reading Mastery* on the Riverside Publishing Off Grade Reading Proficiency test for the grade 4 cohort, and statistically significant negative effects on the State Reading Proficiency test for the grade 5 cohort. According to WWC calculations, the effects were not statistically significant, and the average effect across two cohorts was not large enough to be considered substantively important.[10] Thus, for the comprehension domain, one study showed indeterminate effects.

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).

8. 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 *Reading Mastery* is in Appendix A6.
9. 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 Stockard (2010), no corrections for multiple comparisons were needed. For the *Reading Mastery* studies summarized here, corrections for clustering were needed, so the significance levels may differ from those reported in the original studies.
10. The WWC computes an average effect size as a simple average of the effect sizes across all individual findings within the study domain.

The WWC found *Reading Mastery* to have potentially positive effects on reading fluency and no discernible effects on 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 improvement index for reading fluency is +19 percentile points for a single finding from one study. The average improvement index for comprehension is –7 percentile points, with a range of –20 to +7 percentile points across findings from one study.

Summary

The WWC reviewed 175 studies on *Reading Mastery* for adolescent learners. One of these studies meets WWC evidence standards, and one study meets WWC evidence standards with reservations; the remaining 173 studies do not meet either WWC evidence standards or eligibility screens. Based on these studies, the WWC found potentially positive effects on reading fluency and no discernible effects on comprehension for adolescent learners. The conclusions presented in this report may change as new research emerges.

References

Meets WWC evidence standards

Stockard, J. (2010). *Fourth graders' growth in reading fluency: A pretest-posttest randomized control study comparing* Reading Mastery *and Scott Foresman Basal Reading Program*. Eugene, OR: National Institute for Direct Instruction.

Meets WWC evidence standards with reservations

Yu, L., & Rachor, R. (2000, April). *The two-year evaluation of the three-year* Direct Instruction *program, in an urban public school system.* Presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

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

Adams, G., & Carnine, D. (2003). *Direct Instruction*. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 403–416). New York: Guilford Press. 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.

Adams, G. L., & Englemann, S. (1996). *Research on* Direct Instruction: *25 years beyond* DISTAR. Seattle, WA: Educational Achievement Systems. 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.

Altwerger, B., Arya, P., Jin, L., Jordan, N. L., Laster, B., Martens, P., … Wiltz, N. (2004). When research and mandates collide: The challenges and dilemmas of teacher education in the era of NCLB. *English Education, 36*(2), 119–133. 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.

Altwerger, B., Arya, P., Jin, L., Lang, D., Laster, B., Martens, P., … Wilson, G. P. (2001, December). *The impact of reading programs on second graders' reading*. Symposium conducted at the annual meeting of the National Reading Conference, Miami, FL. 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.

References *(continued)*

American Federation of Teachers. (1999). *Building on the best, learning from what works: Five promising remedial reading intervention programs*. Washington, DC: Author. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Additional source:

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

American Institutes for Research. (1999). *An educators' guide to schoolwide reform*. 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.

Ashworth, D. R. (1999). Effects of *Direct Instruction* and basal reading instruction programs on the reading achievement of second graders. *Reading Improvement, 36*(4), 150–156. 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.

Banta, K. (2002). Direct Instruction Reading Mastery: *A small scale study focused on male students with specific learning disabilities. Does it really work?* Unpublished master's thesis, Cardinal Stritch University, Milwaukee, WI. 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.

Bateman, B. (1991). Teaching word recognition to slow-learning children. *Reading, Writing, and Learning Disabilities, 7*, 1–16. 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.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2002). *Comprehensive school reform and student achievement: A meta-analysis*. Baltimore, MD: CRESPAR/Johns Hopkins University. 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.

Brandes, J. A. (2004). Literacy instruction for students with autism spectrum disorder (ASD): A national survey of special educators (Doctoral dissertation, University of Oklahoma, 2004). *Dissertation Abstracts International, 65*(07A), 230–2558. 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.

Brattland, E. M. (1991). *Oral strategies to develop vocabulary in* Reading Mastery III. Unpublished master's project, Weber State University, Ogden, UT. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Brent, G., & Diobilda, N. (1993). Effects of curriculum alignment versus direct instruction on urban children. *Journal of Educational Research, 86*, 333–338. 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.

Briggs, K. L., & Clark, C. (1997). *Reading programs for students in the lower elementary grades: What does the research say?* Austin, TX: Texas Center for Educational Research. 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.

Brooks, G. (1999). What works for slow readers? *Support for Learning, 14*(1), 27. 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.

Butler, M. T. (2001). Comparison of the effects of direct instruction and basal instruction on the reading achievement of first-grade students identified as students with reading difficulties (Doctoral dissertation, University of Alabama, 2001). *Dissertation Abstracts International, 62*(09A), 203–3002. The study is ineligible for review because it does not use a sample aligned

References *(continued)*

with the protocol—the sample is not within the specified age or grade range.

Butler, P. A. (2003). Achievement outcomes in Baltimore city schools. *Journal of Education for Students Placed at Risk, 8*(1), 33–60. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Calinescu, C. (1997). Outcomes of educational intervention with students with neurological disorders (Master's thesis, University of Toronto [Canada]). *Masters Abstracts International, 36*(06), 75–1440. 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.

Carlson, C. D., & Francis, D. J. (2002). Increasing the reading achievement of at-risk children through *Direct Instruction*: Evaluation of the Rodeo Institute for Teacher Excellence (RITE). *Journal of Direct Instruction, 3*(1), 29–50. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Cheung, A., & Slavin, R. E. (2005). Effective reading programs for English language learners and other language-minority students. *Bilingual Research Journal, 29*(2), 241–267. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Cole, K. N., & Dale, P. S. (1986). Direct language instruction and interactive language instruction with language delayed preschool children: A comparison study. *Journal of Speech and Hearing Research, 29*(2), 206–217. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Cooke, N. L., Gibbs, S. L., Campbell, M. L., & Shalvis, S. L. (2004). A comparison of *Reading Mastery Fast Cycle* and *Horizons Fast Track A-B* on the reading achievement of students with mild disabilities. *Journal of Direct Instruction, 4*(2), 139–151. 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.

Cross, R. W., Rebarber, T., & Wilson, S. F. (2002). Student gains in a privately managed network of charter schools using *Direct Instruction. Journal of Direct Instruction, 2*(1), 3–21. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Crowe, E. C., Connor, C. M., & Petscher, Y. (2009). Examining the core: Relations among reading curricula, poverty, and first through third grade reading achievement. *Journal of School Psychology, 47*(3), 187–214. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Dale, P. S., & Cole, K. N. (1988). Comparison of academic and cognitive programs for young handicapped children. *Exceptional Children, 54*(5), 439–447. 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 sources:

Cole, K. N., Dale, P. S., & Mills, P. E. (1991). Individual differences in language delayed children's responses to direct and interactive preschool instruction. *Topics in Early Childhood Special Education, 11*(1), 99–124.

Cole, K. N., Dale, P. S., Mills, P. E., & Jenkins, J. R. (1993). Interaction between early intervention curricula and student characteristics. *Exceptional Children, 60*(1), 17–28.

Cole, K. N., Mills, P. E., & Dale, P. S. (1989). A comparison of the effects of academic and cognitive curricula for young handicapped children one and two years post program. *Topics in Early Childhood Special Education, 9*(3), 110–127.

Cole, K. N., Mills, P. E., Jenkins, J. R., & Dale, P. S. (2005). Early intervention curricula and subsequent adolescent social development: A longitudinal examination. *Journal of Early Intervention, 27*(2), 71–82.

References *(continued)*

Dale, P. S., Jenkins, J. R., Mills, P. E., & Cole, K. N. (2005). Follow-up of children from academic and cognitive preschool curricula at 12 and 16. *Exceptional Children, 71*(3), 301–317.

Dale, P. S., Mills, P. E., Cole, K. N., & Jenkins, J. R. (2004). When paths diverge: "Errors of prediction" from preschool test scores to later cognitive and academic measures. The *Journal of Special Education, 37*(4), 237–248.

Jenkins, J. R., Dale, P. S., Mills, P. E., Cole, K. N., Pious, C., & Ronk, J. (2006). How special education preschool graduates finish: Status at 19 years of age. *American Educational Research Journal, 43*(4), 737–781.

Mills, P. E., Cole, K. N., Jenkins, J. R., & Dale, P. S. (2002). Early exposure to *Direct Instruction* and subsequent juvenile delinquency: A prospective examination. *Exceptional Children, 69*(1), 85–96.

Mills, P. E., Dale, P. S., Cole, K. N., & Jenkins, J. R. (1995). Follow-up of children from academic and cognitive preschool curricula at age 9. *Exceptional Children, 61*(4), 378–393.

Darch, C., Gersten, R., & Taylor, R. (1987). Evaluation of the Williamsburg County *Direct Instruction* program: Factors leading to success in rural elementary programs. *Research in Rural Education, 4*(3), 111–118. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Deyo, L. L. (2006). *No Child Left Behind according to script: The efficacy of* Reading Mastery. Unpublished master's thesis, New Mexico Highlands University, Las Vegas, NM. 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.

Dowdell, T. (1996). The effectiveness of *Direct Instruction* on the reading achievement of sixth graders (ERIC Reproduction Service no. ED396268). 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.

Dudley-Marling, C., & Paugh, P. (2005). The rich get richer; the poor get *Direct Instruction*. In B. Altwerger (Ed.), *Reading for profit: How the bottom line leaves kids behind* (pp. 156–171). Portsmouth, NH: Heinemann. 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.

Duran, E. (2002). *Systematic instruction in reading for Spanish-speaking students*. Springfield, IL: Charles C. Thomas. 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.

Durnin, S. L. (2008). *The* Reading Mastery *program: A deeper insight into students' opinions regarding reading*. Unpublished master's thesis, Western Oregon University, Monmouth, OR. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Earheart, L. S. (2002). The efficacy of the SRA reading program for disabled learners as measured by the Terra Nova achievement test (Doctoral dissertation, Tennessee State University, 2002). *Dissertation Abstracts International, 63*(08A), 57–2823. 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.

Education Commission of the States. (1999). Direct Instruction. Denver, CO: 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.

Additional source:

Education Commission of the States. (1990). Direct Instruction. Denver, CO: Author.

Ellis, L. A. (2005). *Balancing approaches: Revisiting the educational psychology research on teaching students with learning difficulties*. Victoria, Australia: Australian Council for Educational Research Press. 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.

References *(continued)*

Francis, B. J. (1991). Matching reading programs to students' needs: An examination of alternate programming using a direct instruction program in the regular classroom (Master's thesis, Simon Fraser University [Canada]). *Masters Abstracts International*, *31*(01), 144–61. 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.

Frankhauser, M. A., Tso, M. E., & Martella, R. C. (2001). A comparison of curriculum-specified reading checkout timings and daily 1-minute timings on student performance in *Reading Mastery. Journal of Direct Instruction, 1*(2), 85–96. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Fredrick, L. D., Keel, M. C., & Neel, J. H. (2002). Making the most of instructional time: Teaching reading at an accelerated rate to students at risk. *Journal of Direct Instruction, 2*(1), 57–63. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Funderburk, S. F. (2005). *A research-based, reading approach for severely emotionally disturbed children*. Unpublished doctoral dissertation, University of Georgia, Athens, GA. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Gersten, R., Keating, T., & Becker, W. (1988). The continued impact of the *Direct Instruction* model: Longitudinal studies of follow through students. *Education and Treatment of Children, 11*(4), 318–327. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Gervase, S. J. (2005). Reading Mastery*: A descriptive study of teachers' attitudes and perceptions towards* Direct Instruction. Unpublished master's thesis, Bowling Green State University, Bowling Green, OH. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Glang, A., Ylvisaker, M., Stein, M., Ehlhardt, L., Todis, B., & Tyler, J. (2008). Validated instructional practices: Application to students with traumatic brain injury. *Journal of Head Trauma Rehabilitation, 23*(4), 243–251. 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.

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. *The Journal of Special Education, 39*(2), 66–85.

Hammond, L. (1998). *An examination of four reading programs:* Letterland, *the* Spalding *method (incorporating the* Sound Way Program), Direct Instruction Reading, THRASS (Teaching Handwriting, Reading And Spelling Skills) *to teach beginning literacy in the Kimberley education district*. Broome, W.A., Australia: Kimberley District Education Office. 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.

Harrington, C. J. (1977). *Effectiveness of the nonsense words in the* Reading Mastery *program*. Unpublished master's thesis, Utah State University, Logan, UT. The study is ineligible for

References *(continued)*

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

Herb, M. H. (2005). *The effects of* Reading Mastery *for students with learning disabilities*. Unpublished master's thesis, Pennsylvania State University, Philadelphia, PA. 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.

Herrera, J. A., Logan, C. H., Cooker, P. G., Morris, D. P., & Lyman, D. E. (1997). Phonological awareness and phonetic-graphic conversion: A study of the effects of two intervention paradigms with learning disabled children. Learning disability or learning difference? *Reading Improvement, 34*(2), 71–89. 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.

Herrington, M. S. (1999). A comparative analysis of the *Reading Mastery* and *Silver Burdett Reading* programs for elementary students (Doctoral dissertation, Mississippi State University, 1999). *Dissertation Abstracts International*, *60*(05A), 98–1408. 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.

Hill, K. M. (2004). Improving first-grade reading outcomes: An analysis of a school district reading accountability system (Doctoral dissertation, University of Cincinnati, 2004). *Dissertation Abstracts International*, *65*(08A), 127–2938. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Hogue, C. (2004). *Use of traditional and* Reading Mastery *instruction on students with behavioral and emotional disabilities*. Unpublished master's thesis, Governors State University, University Park, IL. 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.

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 design or a single-case design.

Johnson, S. (1985). *The effects of using the* Reading Mastery Direct Instruction *program with average and above-average kindergarteners: A pilot study*. Unpublished master's thesis, University of Washington, Seattle, WA. 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.

Jones, C. D. (2002). Effects of direct instruction programs on the phonemic awareness abilities of kindergarten students (Doctoral dissertation, University of Virginia, 2002). *Dissertation Abstracts International*, *63*(03), 902A. 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.

Jordan, N. L. (2005). Basal readers and reading as socialization: What are children learning? *Language Arts, 82*(3), 204–213. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Joseph, B. L. (2004). Teacher expectations of low-SES preschool and elementary children: Implications of a research-validated instructional intervention for curriculum policy and school reform (Doctoral dissertation, East Carolina University, 2004). *Dissertation Abstracts International*, *65*(01A), 154–35. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Kamps, D. M., & Greenwood, C. R. (2005). Formulating secondary-level reading interventions. *Journal of Learning Disabilities, 38*(6), 500–509. 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.

Kamps, D. M., Wills, H. P., Greenwood, C. R., Thorne, S., Lazo, J. F., Crockett, J. L., … Swaggart, B. L. (2003). Curriculum influences on growth in early reading fluency for students with academic and behavioral risks: A descriptive study. *Journal of Emotional & Behavioral Disorders, 11*(4), 211–224. The study is

References *(continued)*

ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Additional source:

Kamps, D. M., Wills, H. P., Greenwood, C. R., Thorne, S., Lazo, J. F., Crockett, J. L., … Swaggart, B. L. (2004). Curriculum influences on growth in early reading fluency for students with academic and behavioral risks: A descriptive study. *Journal of Direct Instruction, 4*(2), 189–210.

Kamps, D., Abbott, M., Greenwood, C., ArreagaMayer, C., Wills, H., Longstaff, J., … Walton, C. (2007). Use of evidence-based, small-group reading instruction for English language learners in elementary grades: Secondary-tier intervention. *Learning Disability Quarterly, 30*(3), 153–168. 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.

Kamps, D., Abbott, M., Greenwood, C., Wills, H., Veerkamp, M., & Kaufman, J. (2008). Effects of small-group reading instruction and curriculum differences for students most at risk in kindergarten. *Journal of Learning Disabilities, 41*(2), 101–114. 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.

Keel, M. C., Federick, 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 design or a single-case design.

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.

Knutson, J. S. (2005). The effect of corrective feedback and individualized practice guided by formative evaluation on the reading performance of children who have not made adequate progress in early reading instruction (Doctoral dissertation, University of Oregon, 2005). *Dissertation Abstracts International*, *66*(07A), 126–2531. 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.

Kuder, S. J. (1990). Effectiveness of the *DISTAR* reading program for children with learning disabilities. *Journal of Learning Disabilities, 23*(1), 69–71. 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.

Kuder, S. J. (1991). Language abilities and progress in a *Direct Instruction* reading program for students with learning disabilities. J*ournal of Learning Disabilities, 24*(2), 122–127. 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.

Leach, D., & Siddall, S. (1990). Parental involvement in the teaching of reading: A comparison of hearing reading, paired reading, pause, prompt, praise, and direct instruction methods. *British Journal of Educational Psychology, 60*, 349–355. 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.

League, M. B. (2001). The effect of the intensity of phonological awareness instruction on the acquisition of literacy skills (Doctoral dissertation, University of Florida, 2001). *Dissertation Abstracts International*, *62*(10), 3299A. 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.

LeCapitaine, K. J. (2002). *Does the use of a* Direct Instruction *reading program affect the early literacy skills in an urban five-year-old kindergarten class?* Unpublished master's thesis, Cardinal Stritch University, Milwaukee, WI. 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.

References *(continued)*

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.

Additional sources:

Ligas, M. R., & Vaughan, D. W. (1999). *Alliance of Quality Schools: 1998–99 evaluation report*. Broward, FL: Broward County Schools.

Varela-Russo, C., Blasik, K. A., & Ligas, M. R. (1998). *Alliance of Quality Schools evaluation report*. Ft. Lauderdale, FL: School Board of Broward County.

Lovett, M., Palma, M., Frijters, J., Steinbach, K., Temple, M., Benson, N., & Lacerenza, L. (2008). Interventions for reading difficulties: A comparison of response to intervention by ELL and EFL struggling readers. *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.

Lutz, A. R. (2004). *The effectiveness of the* Reading Mastery *reading program when teaching learning support students how to read*. Unpublished master's thesis, Gratz College, Melrose Park, 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.

Mac Iver, M. A., & Kemper, E. (2002). Guest editors' introduction: Research on *Direct Instruction* in reading. *Journal of Education for Students Placed at Risk, 7*(2), 107–116. 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.

Mac Iver, M. A., & Kemper, E. (2002). The impact of *Direct Instruction* on elementary students' reading achievement in an urban school district. *Journal of Education for Students Placed at Risk, 7*(2), 197–220. 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.

Additional source:

Mac Iver, M. A., Kemper, E., & Stringfield, S. (2003). *The Baltimore curriculum project: Final report of the four-year evaluation study*. Baltimore, MD: Center for Research on the Education of Students Placed at Risk.

Marasch, K. L. (1976). *A formative evaluation of the* Reading Mastery *program*. Unpublished master's thesis, Utah State University, Logan, UT. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Marchand-Martella, N. E., Martella, R. C., Kolts, R. L., Mitchell, D., & Mitchell, C. (2006). Effects of a three-tier strategic model of intensifying instruction using a research-based core reading program in grades K–3. *Journal of Direct Instruction, 6*(1), 49–72. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Marston, D., Deno, S. L., Kim, D., & Diment, K. (1995). Comparison of reading intervention approaches for students with mild disabilities. *Exceptional Children, 62*(1), 20–37. 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.

McGraw-Hill Education. (2002). Eshelman Avenue Elementary, Lomita, California. In *Results with* Reading Mastery (pp. 16–17). New York: McGraw-Hill. The study is ineligible for review because it does not use a comparison group design or a single-case design.

McGraw-Hill Education. (2002). Fort Worth Independent School District, Fort Worth, Texas. In *Results with* Reading Mastery (pp. 4–5). New York: McGraw-Hill. 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.

McGraw-Hill Education. (2002). Lebanon School District, Lebanon, Pennsylvania. In *Results with* Reading Mastery (pp. 8–9). New York: McGraw-Hill. The study is ineligible for review because it does not use a comparison group design or a single-case design.

References *(continued)*

McGraw-Hill Education. (2002). Park Forest–Chicago Heights School District 163, Chicago, Illinois. In *Results with* Reading Mastery (pp. 10–11). New York: McGraw-Hill. The study is ineligible for review because it does not use a comparison group design or a single-case design.

McGraw-Hill Education. (2002). Portland Elementary School, Portland, Arkansas. In *Results with* Reading Mastery (pp. 2–3). New York: McGraw-Hill. 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.

McGraw-Hill Education. (2002). Roland Park Elementary/Middle School, Baltimore, Maryland. In *Results with* Reading Mastery (pp. 12–13). New York: McGraw-Hill. The study is ineligible for review because it does not use a comparison group design or a single-case design.

McGraw-Hill Education. (2002). Wilson Primary School, Phoenix, Arizona. In R*esults with* Reading Mastery (pp. 6–7). New York: McGraw-Hill. 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.

McIntyre, E., Rightmyer, E. C., & Petrosko, J. P. (2008). Scripted and non-scripted reading instructional models: Effects on the phonics and reading achievement of first-grade struggling readers. *Reading & Writing Quarterly, 24*(4), 377–407. 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.

Mitchell, D. R. (2008). *What really works in special and inclusive education: Using evidence-based teaching strategies*. New York: Routledge. 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.

Morgan, R. L., Menlove, R., Salzberg, C. L., & Hudson, P. (1994). Effects of peer coaching on the acquisition of direct instruction skills by low-performing preservice teachers. *The Journal of Special Education, 28*(1), 59–76. The study is ineligible for review because it does not include a student outcome.

Morgenstern, B. D. (2002). A comparison of high- and low-frequency criteria on reading agility, retention, endurance and *Direct Instruction Reading Mastery* checkout performance of elementary students academically at-risk (Doctoral dissertation, The Ohio State University, 2002). *Dissertation Abstracts International*, *63*(07A), 86–2464. 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.

Mosley, A. M. (1997). *The effectiveness of* Direct Instruction *on reading achievement*. (ERIC Reproduction Service no. ED402553). 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.

Nanda, A. O., & Fredrick, L. D. (2007). The effects of combining repeated reading with *Reading Mastery* on first graders' oral reading fluency. *Journal of Direct Instruction, 7*(1), 17–27. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Nealy-Maxwell, M. A. (2004). *Investigating how the use of* Direct Instruction *can complement a balanced literacy program related to reading comprehension and decoding skills*. Unpublished master's thesis, Mount Mary College, Milwaukee, WI. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Neely, M. (1995). The multiple effects of whole language, precision teaching, and *Direct Instruction* on first-grade story-reading. *Effective School Practices, 14*(4), 33–42. 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.

Nelson, L. H. (2008). *Effective strategies a first grade teacher in a* Direct Instruction *setting can use to improve reading fluency in beginning readers*. Unpublished master's thesis, Hamline University, Saint Paul, MN. The study is ineligible for review because it

References *(continued)*

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

O'Brien, D. M., & Ware, A. M. (2002). Implementing research-based reading programs in the Fort Worth Independent School District. *Journal of Education for Students Placed at Risk, 7*(2), 167–195. 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.

Ocokoljich, E. D. (1997). *The effects of* Reading Mastery I *and* II *on the reading achievement of first and second grade students identified as having low phonological awareness skills.* Unpublished master's thesis, University of Wisconsin–Madison, Madison, WI. The study is ineligible for review because it does not use a comparison group design or a single-case design.

O'Connor, R. E., Jenkins, J. R., Cole, K. N., & Mills, P. E. (1993). Two approaches to reading instruction with children with disabilities: Does program design make a difference? *Exceptional Children, 59*(4), 312–323. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Additional source:

O'Connor, R. E., Jenkins, J. R., Cole, K. N., & Mills, P. E. (1992, April). *Two approaches to reading instruction for children with disabilities: Does program design make a difference?* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Paired Learning Project. (1989). Paired learning: Tutoring by non-teachers. *Paired Reading Bulletin, 5*. 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.

Prager, A. J. (2008). A comparison of linear versus spiral multiple exemplar instruction on derived abstracted textual responses of preschool children (Doctoral dissertation, Columbia University, 2008). *Dissertation Abstracts International*, *68*(11A), 145–4651. 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.

Richardson, E. (1978). An assessment of two methods for remediating reading deficiencies. *Reading Improvement, 15*(2), 82–95. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

Riepl, J., Marchand-Martella, N., & Martella, R. (2008). The effects of *Reading Mastery Plus* on the beginning reading skills of students with intellectual and developmental disabilities. *Journal of Direct Instruction, 8*(1), 29–39. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Rodman, M. L. (2008). A study of intensive, systematic *Direct Instruction* for an autistic child (autism). *Dissertation Abstracts International*, 68(7-A), 2896. 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.

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

Rothenbusch, T. (1999). *The effects of a precision teaching/* Direct Instruction *reading program on the reading achievement of elementary school students*. Unpublished master's thesis, University of Northern British Columbia, Prince George, BC, Canada. 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.

Ryder, R. J., Burton, J. L., & Silberg, A. (2006). Longitudinal study of *Direct Instruction* effects from first through third grades. *Journal of Educational Research, 99*(3), 180–191. The study is

References *(continued)*

ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Additional source:

Ryder, R. J., Sekulski, J. L., & Silberg, A. (2003). *Results of* Direct Instruction *reading program evaluation. Longitudinal results: First through third grade 2000–2003*. Unpublished manuscript.

Schieffer, C., Marchand-Martella, N., Martella, R., & Simonsen, F. (2002). *The research base for* Reading Mastery. 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.

Schieffer, C., Marchand-Martella, N. E., Martella, R. C., Simonsen, F. L., & Waldron-Soler, K. M. (2002). An analysis of the *Reading Mastery* program: Effective components and research review. *Journal of Direct Instruction, 2*(2), 87–119. 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.

Schleider, M. J. (2006). Predicting responsiveness-to-intervention in reading from curriculum-based, social skill and problem behavior measurements (Doctoral dissertation, University of North Carolina at Charlotte, 2006). *Dissertation Abstracts International*, *68*(01A), 130–151. 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.

Schmidt, C. J. (2008). *The effects of Reading Recovery*, Direct Instruction, *and Literacy Collaborative on 2nd graders' reading*. Unpublished master's thesis, Carthage College, Kenosha, WI. 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.

Schug, M. C., Tarver, S. G., & Western, R. D. (2001). *Direct Instruction* and the teaching of early reading: Wisconsin's teacher-led insurgency. *Wisconsin Policy Research Institute Report, 14*(2). Hartland, WI: Wisconsin Policy Research Institute, Inc. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Sexton, C. W. (1989). Effectiveness of the *DISTAR Reading I* program in developing first graders' language skills. *Journal of Educational Research, 82*(5), 289–293. 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.

Shelton, N. R. (2005). First do no harm: Teachers' reactions to mandated *Reading Mastery*. In B. Altwerger (Ed.), *Reading for profit: How the bottom line leaves kids behind* (pp. 184–198). Portsmouth, NH: Heinemann. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Shippen, M. E., Houchins, D. E., Calhoon, M. B., Furlow, C. F., & Sartor, D. L. (2006). The effects of comprehensive school reform models in reading for urban middle school students with disabilities. *Remedial and Special Education, 27*(6), 322–328. 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.

Simmons, D. C., Coyne, M. D., Kwok, O., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention. *Journal of Learning Disabilities, 41*(2), 158–173. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Slavin, R. E., & 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., Lake, C., Davis, S., & Madden, N. A. (2009). *Effective programs for struggling readers: A best-evidence synthesis*. Baltimore, MD: Johns Hopkins University, Center

References *(continued)*

for Data-Driven Reform 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.

Smith, S. L. B. (2002). A cross-case study of teacher perceptions of program design and teaching efficacy: The Seaview Reading Support Program (California) (Doctoral dissertation, University of San Diego, 2002). *Dissertation Abstracts International*, *63*(01A), 246–131. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Smolkowski, K., Biglan, A., Barrera, M., Taylor, T., Black, C., & Blair, J. (2005). *Schools and Homes in Partnership (SHIP)*: Long-term effects of a preventive intervention focused on social behavior and reading skill in early elementary school. *Prevention Science: The Official Journal of the Society for Prevention Research, 6*(2), 113–125. 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.

Snider, V. E. (1990). *Direct Instruction* with average first-graders. *Reading Improvement, 27*, 143–148. 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.

Sprinkman, A. (2001). *A comparison of reading achievement made by LD and low IQ students using a* Direct Instruction *reading program*. Unpublished master's thesis, Cardinal Stritch University, Milwaukee, WI. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2003). *City Springs Elementary School, Baltimore, Maryland*. DeSoto, TX: Author. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). *All grade 3 students in two Monroe, Wisconsin elementary schools score proficient or advanced in reading*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/monroe_di.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). *Barren County Elementary Schools post highest reading scores ever*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/barren_di1.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). *California Blue Ribbon school closes achievement gap with* Reading Mastery. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/richfield_di.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). *Delaware charter school students maintain high reading scores*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/east_side_di.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). *Florida elementary students master reading in preparation for junior high*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/clay_hill_di.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). *Miami elementary school boosts FCAT scores with* Reading Mastery. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/parkway_di.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). *Milwaukee elementary nearly doubles reading scores*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/honey_creek_di1.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

References *(continued)*

SRA/McGraw-Hill. (2005). *Oregon Reading First project uses* Reading Mastery Plus *as core reading program*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/MiltonFreewater_DI.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). *Phoenix inner-city students strive toward national reading average*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/wilson_di.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). Reading Mastery *helps Florida students advance two grade levels in reading*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/gulf_di.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). Reading Mastery Plus *helps Colorado school achieve AYP for first time*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/ivywild_di.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2005). *Washington elementary students excel on WASL, ITBS with* Reading Mastery Plus. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/evergreen_di.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2006). *Cleveland school keeps* Reading Mastery *as curriculum core*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/louisa_may_alcott_di1.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2006). *DIBELS scores advance to grade level with* Reading Mastery. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/Edgewood_DI.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2006). *Exceptional education and regular education students excel with* Direct Instruction. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/Iredell_DI_FNL.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2006). *Florida school moves from D grade to A with* Reading Mastery. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/AltaVista_DI.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2006). *Native American school uses Reading First grant to implement* Direct Instruction. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/NayAhShing_DI.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2006). Reading Mastery, Corrective Reading *help students with disabilities achieve significant academic growth*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/Clover_DI.pdf. 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. (2006). *Reading proficiency more than doubles among Putnam County special education students*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/PutnamCo.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2006). *Struggling Milwaukee readers make strong gains with* Direct Instruction. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/EastHS_DI.pdf. The study is ineligible for

References *(continued)*

review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2006). *Utah school district maintains high language arts scores with* Direct Instruction. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/Cache.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2007). *Low-performing Kentucky school on its way to high-performing with* Reading Mastery. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/Highland_DI_FNL.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2007). Reading Mastery *helps special education students meet state reading standards*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/Marmarton_DI.pdf. 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. (2007). *Reading scores rise at Alabama elementary with* Reading Mastery Plus. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/ElbaElem_DI.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2007). *SRA/McGraw-Hill's reading programs bring increases in Baltimore's scores*. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/Baltimore_DI_07.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2007). *Title I schools in North Carolina district meet all-state reading targets with* Direct Instruction. Retrieved from SRA website: https://www.sraonline.com/download/DI/EfficacyReports/Brunswick-ER_FNL.pdf. The study is ineligible for review because it does not use a comparison group design or a single-case design.

SRA/McGraw-Hill. (2009). *A report on the effects of SRA/McGraw-Hill's* Reading Mastery, Signature Edition: *A response to intervention solution*. Desoto, TX: SRA/McGraw-Hill. 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.

Stewart, R. M., Benner, G. J., Martella, R. C., & Marchand-Martella, N. E. (2007). Three-tier models of reading and behavior: A research review. *Journal of Positive Behavior Interventions, 9*(4), 239–253. 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.

Stockard, J. (2008). *Improving first grade reading achievement in a large urban district: The effects of NIFDI-supported implementation of* Direct Instruction *in the Baltimore City public school system* (Technical report no. 2008-1). Eugene, OR: National Institute for Direct Instruction. 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.

Stockard, J. (2008). *Reading achievement in a direct instruction school and a "three tier" curriculum school* (Technical report no. 2008-5). Eugene, OR: National Institute for Direct Instruction. 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.

Stockard, J. (2008). *The long-term impact of NIFDI-supported implementation of* Direct Instruction *on reading achievement: An analysis of fifth graders in the Baltimore City public school system*. Eugene, OR: National Institute for Direct Instruction. 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.

Stockard, J. (2009). *Promoting reading achievement and countering the "fourth grade slump": The impact of* Direct

References *(continued)*

Instruction *on reading achievement in fifth grade*. Eugene, OR: National Institute for Direct Instruction. 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.

Stockard, J., & Englemann, K. (2008). *Academic kindergarten and later academic success: The impact of* Direct Instruction (Technical report no. 2008-7). Eugene, OR: National Institute for Direct Instruction. 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.

Sullivan, M. (2002). Reading Mastery *versus word study instruction as it pertains to third graders' reading achievement scores*. Unpublished educational specialist's thesis, Western Kentucky University, Bowling Green, KY. 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.

Thames, D., Kazelskis, R., & Kazelskis, C. R. (2006, November). *Reading performance of elementary students: Results of a five-year longitudinal study of direct reading instruction*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Birmingham, AL. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Thomson, B. (1991). Pilot study of the effectiveness of a *Direct Instruction Model* (*Reading Mastery Fast Cycle*) as a supplement to a literature based delivery model (Houghton-Mifflin *Integrated Reading Program*) in two regular first grade classrooms. *Florida Educational Research Council Research Bulletin, 23*(2), 3–23. 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.

Tobin, K. (2000). *A comparison between* Horizons Fast Track A-B *and Silver, Burdett, and Ginn reading curricula in first grade: June 2000 final report*. Eugene, OR: National Institute for Direct Instruction. 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.

Torgesen, J. K., & King, R. (2000). *Improving the effectiveness of reading instruction in one elementary school: A description of the process* (Technical report no. 3). Tallahassee, FL: Florida Center for Reading Research. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Traweek, D., & Berninger, V. W. (1997). Comparisons of beginning literacy programs: Alternative paths to the same learning outcome. *Learning Disability Quarterly, 20*(2), 160–168. 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.

Trout, A. L., Epstein, M. H., Mickelson, W. T., Nelson, J. R., & Lewis, L. M. (2003). Effects of a reading intervention for kindergarten students at risk for emotional disturbance and reading deficits. *Behavioral Disorders, 28*(3), 313–326. 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.

Umbach, B., Darch, C., & Halpin, G. (1989). Teaching reading to low performing first graders in rural schools: A comparison of two instructional approaches. *Journal of Instructional Psychology, 16*(3), 112–121. 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.

Urdegar, S. M. (1998). *Evaluation of the* Success for All *program, 1997–98*. Miami, FL: Miami-Dade Public Schools, Office of Evaluation Research. 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.

Additional sources:

Urdegar, S. M. (2000). *Evaluation of the* Success for All *program, 1998–99*. Miami, FL: Miami-Dade Public Schools, Office of Evaluation Research.

References *(continued)*

Urdegar, S. M. (2001). *Evaluation of the* Success for All *program, 1999–2000*. Miami, FL: Miami-Dade Public Schools, Office of Evaluation Research.

Valesano, L. (2000). *Advantages of using* Direct Instruction *to teach reading*. Unpublished starred paper, St. Cloud State University, St. Cloud, MN. 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.

Valesano, L. (2000). Reading Mastery: *A phonemic reading program to teach reading*. Unpublished starred paper, St. Cloud State University, St. Cloud, MN. 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 design or a single-case design.

Watkins, T. B. (2008). A comparative analysis of the effectiveness of *Direct Instruction* reading on African American, Caucasian, and Hispanic students (Doctoral dissertation, Delta State University, 2008). *Dissertation Abstracts International*, *69*(03A), 104–923. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Weinstein, G., & Cooke, N. L. (1992). The effects of two repeated reading interventions on generalization of fluency. *Learning Disability Quarterly, 15*(1), 21–28. 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.

Weisberg, P. (1988). *Direct Instruction* in the preschool. *Education & Treatment of Children, 11*(4), 349–363. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

White, W. A. T. (1988). A meta-analysis of the effects of *Direct Instruction* in special education. *Education & Treatment of Children, 11*(4), 364–374. The study is ineligible for review because it does not occur within the time frame specified in the protocol.

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.

Wilson, G. P., Wiltz, N. W., & Lang, D. (2005). The impact of *Reading Mastery* on children's reading strategies. In B. Altwerger (Ed.), *Reading for profit: How the bottom line leaves kids behind* (pp. 172–183). Portsmouth, NH: Heinemann. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Wilson, P., Martens, P., & Arya, P. (2005). Accountability for reading and readers: What the numbers don't tell. *Reading Teacher, 58*(7), 622–631. 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.

Wiltz, N., & Wilson, G. P. (2005). An inquiry into children's reading in one urban school using SRA *Reading Mastery* (*Direct Instruction*). *Journal of Literacy Research, 37*(4), 493–528. 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.

Wrobel, S. (1996). *The effectiveness of* Direct Instruction *on the various reading achievement categories*. (ERIC Reproduction Service no. ED395292). 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.

References *(continued)*

Zayac, R. M. (2008). *Direct Instruction* reading: Effects of the *Reading Mastery Plus* Level K curriculum on preschool children with developmental delays (Doctoral dissertation, Auburn University, 2008). *Dissertation Abstracts International*, *69*(10B), 226–6458. 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.

Ziech, A. (2005). *Improving the literacy skills of a struggling reader through* Direct Instruction *and guided reading*. Unpublished master's thesis, Cardinal Stritch University, Milwaukee, WI. 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 A1.1 Study characteristics: Stockard, 2010

Characteristic	Description
Study citation	Stockard, J. (2010). *Fourth graders' growth in reading fluency: A pretest-posttest randomized control study comparing* Reading Mastery *and* Scott Foresman Basal Reading Program. Eugene, OR: National Institute for Direct Instruction.
Participants	In this randomized study, 58 general education elementary students were assigned to a treatment or control condition, using alternative assignment with a random start technique. The author used a class list (arranged in alphabetical order) to conduct the assignment. First, the author used a random numbers table to determine where to begin in the class list. Second, the author used a coin flip to determine whether the assignment would start with the treatment or control group. Finally, after this initial assignment was determined, the author assigned each additional student in an alternating fashion to the treatment or control group. For example, in the first step, the author might have started the assignment with John Smith (based on a random numbers table). In the second step, the author might have assigned John Smith to the treatment group (based on the coin flip). Then, the author would have proceeded through the rest of the ordered class list (alternating between the control and treatment groups). Four classrooms participated in the study. Two pairs of teachers were formed and then—within these pairs—teachers were randomly assigned to the treatment or control group via a coin flip.[1] Students were predominantly non-Hispanic whites from middle-income families. The analysis sample consisted of 29 fourth-grade students who received *Reading Mastery* and 28 fourth-grade students in the comparison group.
Setting	The study was conducted in a midwestern elementary school.
Intervention	Beginning in the fall of 2009, students in the treatment condition received instruction for 90 minutes a day in the SRA/McGraw-Hill program, *Reading Mastery Signature Edition.* Students were exposed to *Reading Mastery* over five months.
Comparison	The control group received instruction for 90 minutes a day in the *Scott Foresman Basal Reading Program*, which the school had been using in prior years.
Primary outcomes and measurement	Data on the AIMS Web Curriculum-Based Measurement Words Read Correct were gathered in the spring of 2009 before instruction began (for use as a baseline measure), in the fall of 2009 shortly after the start of the school year, and again in winter of 2010, approximately halfway through the school year. For a more detailed description of this outcome measure, see Appendix A2.1.
Staff or teacher training	No information about training was provided.

1. The study author did not describe how students assigned to the treatment group were assigned to the two treatment teachers or how students assigned to the control group were assigned to the two control teachers.

Appendix A1.2 Study characteristics: Yu & Rachor, 2000

Characteristic	Description
Study citation	Yu, L., & Rachor, R. (2000, April). *The two-year evaluation of the three-year* Direct Instruction *program, in an urban public school system*. Presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
Participants	This retrospective quasi-experiment included students in three elementary schools that participated in *Reading Mastery* for two consecutive years. Each of the three *Reading Mastery* schools was matched with a school with a similar level of poverty and percentage of minority students. Then, *Reading Mastery* students were matched with comparison students in the same grade on the basis of race, gender, free lunch status, and reading achievement test scores. This review focuses on findings from students who were in grades 4 and 5[1] during the first year of program implementation (1997/98). The analysis sample for students who were in grade 4 in 1997/98 consisted of 71 students who received *Reading Mastery* and 71 matched comparison students. The analysis sample for students who were in grade 5 in 1997/98 consisted of 81 students in the *Reading Mastery* group and 81 matched comparison students. Of the students, more than 96% were African-American, and about 80% were eligible for free or reduced-price lunch. The study reported students' outcomes after two years of program implementation; these findings were used in the intervention ratings and can be found in Appendix A3.2. Additional findings reflecting students' outcomes after one year of program implementation can be found in Appendix A4.
Setting	The study was conducted in six elementary schools in a northwestern urban public school system.
Intervention	The complete *Reading Mastery* program was adopted, including all materials and teacher training. The program provided scripted, carefully sequenced lessons; rapid pacing; and responses of the students in unison as well as in individual turns. Students were exposed to *Reading Mastery* over two school years.
Comparison	The comparison group received the standard instruction provided in the regular school curriculum.
Primary outcomes and measurement	For both the pretest and posttest, students took the State Reading Proficiency Test and the Riverside Publishing Off Grade Reading Proficiency Test. The State Reading Proficiency Test provided data at pretest and after two years of intervention implementation for the 5th-grade group and data after one year of intervention implementation for the 4th-grade group. The Riverside Publishing Off Grade Reading Proficiency Test provided data at pretest and after two years of intervention implementation for the 4th-grade group, and data after one year of intervention implementation for the 5th-grade group. For a more detailed description of these outcome measures, see Appendix A2.2.
Staff or teacher training	Teachers participated in training prior to implementation and were provided with ongoing consultations from the provider for the program duration. No additional details about training were provided.

1. Findings for 3rd grade are outside the scope of the Adolescent Literacy review.

Appendix A2.1 Outcome measures for the reading fluency domain

Outcome measure	Description
AIMS Web Curriculum-Based Measurement Words Read Correct	When a child reads graded passages aloud for one minute, this curriculum-based AIMS web assessment gathers data on reading fluency in the form of number of words read correctly. The assessment has well-established validity and reliability. AIMS Web is a benchmark and progress-monitoring system based on direct, frequent, and continuous student assessment (as cited in Stockard, 2010).

Appendix A2.2 Outcome measures for the comprehension domain

Outcome measure	Description
Reading comprehension construct	
Riverside Publishing Off Grade Reading Proficiency Test	This test for 3rd- and 5th-grade students is designed to measure students' reading proficiency. Test items, which are based on fiction, nonfiction, and poetry reading passages, are designed to capture four strands of learning outcomes defined by the publisher as (1) examining meaning given a fiction or poetry text (examining meaning indicates that students are able to comprehend the overall meaning of what they have read), (2) extending meaning given a fiction or poetry text (extending meaning indicates that students can interpret what they have read and infer beyond the text), (3) examining meaning given a nonfiction text, and (4) extending meaning given a nonfiction text (as cited in Yu & Rachor, 2000).
State Reading Proficiency Test	This state annual reading test is administered to 4th- and 6th-grade students (as cited in Yu & Rachor, 2000). No additional information on the test was provided.

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

Outcome measure	Study sample	Sample size (classrooms/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Reading Mastery* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*Reading Mastery* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Stockard, 2010[7,8]								
AIMS Web CBM Words Read Correct	Grade 4	4/57	149.0 (29.5)	134.7 (27.8)	14.30	0.49	ns	+19
Domain average for reading fluency (Stockard, 2010)[9]						**0.49**	**ns**	**+19**

ns = not statistically significant

AIMS Web CBM = AIMS Web Curriculum-Based Measurement

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the reading fluency domain.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect-size calculation, see WWC Procedures and Standards Handbook, Appendix B.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. 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 comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results for the intervention group.
7. The *Reading Mastery* group mean outcome values for Stockard (2010) are the unadjusted control group posttest means plus the difference in mean gains between the intervention and control groups. Control group means are unadjusted.
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 Stockard (2010), a correction for clustering was needed, so the significance levels may differ from those reported in the original study.
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 comprehension domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Reading Mastery* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*Reading Mastery* –comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Yu & Rachor, 2000[7,8]								
Riverside Publishing Off Grade Reading Proficiency Test	Grade 4	6/142	209.01 (20.16)	205.46 (18.42)	3.55	0.18	ns	+7
State Reading Proficiency Test	Grade 5	6/162	201.15 (25.30)	213.80 (23.42)	–12.65	–0.52	ns	–20
Domain average for comprehension (Yu & Rachor, 2000)[9]						**–0.17**	**na**	**–7**

ns = not statistically significant
na = not applicable

1. This appendix reports two-year findings considered for the effectiveness rating and the average improvement indices for the comprehension domain. One-year findings from Yu and Rachor (2000) are not included in these ratings but are reported in Appendix A4.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect-size calculation, see WWC Procedures and Standards Handbook, Appendix B.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. 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 comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results for the intervention group.
7. The *Reading Mastery* group mean outcome values for Yu and Rachor (2000) differ from those presented in the paper. The WWC calculated the program group mean using a difference-in-differences approach (see WWC Handbook)—calculating the program means by adding the impact of the program (i.e., difference in mean gains between the intervention and control groups) to the unadjusted control group posttest means.
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 Yu and Rachor (2000), a correction for clustering was needed, so the significance levels may differ from those reported in the original study.
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 Summary of one-year implementation findings for the comprehension domain[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Reading Mastery* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*Reading Mastery* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Yu & Rachor, 2000[7,8]								
State Reading Proficiency Test	Grade 4	6/142	207.13 (12.28)	200.92 (12.18)	6.21	0.51	ns	+19
Riverside Publishing Off Grade Reading Proficiency Test	Grade 5	6/162	207.42 (21.33)	204.95 (24.86)	2.47	0.11	ns	+4

ns = not statistically significant

1. This appendix presents one-year findings for measures that fall in the comprehension domain. Two-year findings were used for rating purposes and are presented in Appendix A3.2.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect-size calculation, see WWC Procedures and Standards Handbook, Appendix B.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. 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 comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
7. The *Reading Mastery* group and control group mean outcome values for Yu and Rachor (2000) are the unadjusted one-year posttest means.
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 Yu and Rachor (2000), a correction for clustering was needed, so the significance levels may differ from those reported in the original study.

Appendix A5.1 *Reading Mastery* 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 reading fluency outcome domain, the WWC rated *Reading Mastery* as having potentially positive effects for adolescent learners. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, or negative effects) were not considered, as *Reading Mastery* was assigned the highest applicable rating.

Rating received

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.

  Met. One study showed substantively important positive effects.

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.

  Met. No studies showed statistically significant or substantively important negative effects. No studies showed indeterminate effects.

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 statistically significant positive effects.

AND

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

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

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. For a complete description, see the WWC Procedures and Standards Handbook, Appendix E.

Appendix A5.2 *Reading Mastery* 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 comprehension outcome domain, the WWC rated *Reading Mastery* 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.

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.

(continued)

Appendix A5.2 *Reading Mastery* 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.

Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence[1]
		Schools	Students	
Alphabetics	na	na	na	na
Reading fluency	1	1	57	Small
Comprehension	1	6	304	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.