

What Works Clearinghouse



Beginning Reading

Updated May 2012

Peer-Assisted Learning/Literacy Strategies

Program Description¹

Peer-Assisted Learning Strategies and a similar program known as *Peer-Assisted Literacy Strategies* are peer-tutoring programs that supplement the primary reading curriculum (Fuchs, Fuchs, Kazdan, & Allen, 1999; Mathes & Babyak, 2001). This review uses the acronym *PALS* to encompass both programs and their respective full names when referring to a specific program. Students in *PALS* classrooms work in pairs on reading activities intended to improve reading accuracy, fluency, and comprehension. Students in the pairs—who alternately take on the role of tutor and tutee—read aloud, listen to their partner read, and provide feedback during various structured activities. Teachers train students to use the following learning strategies: passage reading with partners, paragraph “shrinking” (or describing the main idea), and prediction relay (predicting what is likely to happen next in the passage). *PALS* includes separate versions for kindergarten and grade 1. *Peer-Assisted Learning Strategies* also includes versions for grades 2–3 (which are part of a larger set produced for grades 2–6).²

Report Contents	
Overview	p. 1
Program Information	p. 2
Research Summary	p. 4
Effectiveness Summary	p. 5
References	p. 7
Research Details for Each Study	p. 11
Outcomes Measures for Each Domain	p. 16
Findings Included in the Rating for Each Outcome Domain	p. 17
Supplemental Single-Case Design Studies	p. 20
Endnotes	p. 21
Rating Criteria	p. 22
Glossary of Terms	p. 23

Research³

Two studies of *PALS* that fall within the scope of the Beginning Reading review protocol meet What Works Clearinghouse (WWC) evidence standards, and one study meets WWC evidence standards with reservations. These three studies included 3,130 beginning readers in kindergarten and grade 1 in four states. Based on these studies, the WWC considers the extent of evidence for *PALS* on beginning readers to be medium to large for the alphabetic domain and small for the fluency and comprehension domains.

Effectiveness

PALS was found to have potentially positive effects on alphabetic, no discernible effects on fluency, and mixed effects on comprehension for beginning readers.

Table 1. Summary of findings⁴

Outcome domain	Rating of effectiveness	Improvement index (percentile points)		Number of studies	Number of students	Extent of evidence
		Average	Range			
Alphabetic	Potentially positive effects	+14	-17 to +36	2	3,000	Medium to large
Fluency	No discernible effects	-8	-8 to -9	1	41	Small
Comprehension	Mixed effects	+3	-12 to +17	2	171	Small

Program Information

Background

Developed by Lynn and Doug Fuchs in 1997,⁵ *Peer-Assisted Learning Strategies* is distributed by the Vanderbilt Kennedy Center for Research on Human Development. Address: Vanderbilt University, Attn: Flora Murray/PALS Orders, Peabody Box 228, Nashville, TN 37203-5701. Email: flora.murray@vanderbilt.edu. Web: <http://kc.vanderbilt.edu/pals/>. Telephone: (615) 343-4782. Fax: (615) 343-1570.

Patricia Mathes (currently affiliated with Southern Methodist University) is the primary author *Peer-Assisted Literacy Strategies*.⁶ Mathes's *Peer-Assisted Literacy Strategies* is distributed by Sopris. Address: 4185 Salazar Way, Frederick, CO 80504. Email: customerservice@cambiumlearning.com. Web: <http://www.soprislearning.com>. Telephone: (800) 547-6747.

Program details

This report focuses on both *PALS* reading programs for kindergarten and grade 1 and *Peer-Assisted Learning Strategies* for grades 2 and 3, which are part of a set produced for grades 2–6. In each of these versions of the program, students engage in peer-tutoring routines through a series of structured interactions. Teachers assign students to pairs based on an area in which one student is deficient and the other is proficient (initially, the former serves as the tutee and the latter as the tutor). Throughout the intervention, students are assigned different partners and have the opportunity to be both the provider and recipient of tutoring. *PALS* activities last 35 minutes per session and are intended to be implemented three to four times a week. A typical lesson includes the following activities:

1. Partner reading—the reader (or tutee) reads aloud, receiving immediate corrective feedback if words are mispronounced. The program calls for the stronger reader in each pair to read first, which is designed to provide an opportunity for the weaker reader in the pair to preview the passage and review difficult words before it is his or her turn to reread the same text. Students switch roles after five-minute blocks.
2. Paragraph “shrinking”—the reader states the main idea (i.e., who or what the passage is about), gives a 10-word summary of the passage, and provides a sequential retelling of the important events of the passage.
3. Prediction relay—the reader predicts what is likely to happen next on the next page, reads aloud from the page, and summarizes the just-read text, with the tutor deciding whether the predictions are accurate. Students switch roles after five-minute blocks.

K-PALS Reading (Kindergarten PALS) and *First-Grade PALS* include a set of 70 student lesson sheets. Teachers choose appropriate reading material for partner reading. Student reading materials are not provided as part of the grades 2–6 version of *Peer-Assisted Learning Strategies* because teachers are supposed to select appropriate reading materials. A typical lesson for first-grade students begins with 15 minutes of Sounds and Words, which focuses on learning to hear and identify sounds, sounding out words, learning sight words, and practicing passage reading. The next 15 minutes are spent on Story Sharing, which focuses on predicting story plots, oral reading, and retelling stories. A typical lesson for students in grades 2–6 includes specific activities to improve reading accuracy, fluency, and reading comprehension.

The motivational system used for the grades K–6 *PALS* program involves students earning points for their team by reading sentences without error, working hard, and identifying the correct subject and main idea during paragraph summary. Points are awarded by tutors and teachers and are recorded by students on scorecards. *Peer-Assisted Learning Strategies* offers teacher training in an all-day workshop at which teachers learn to implement the program through modeling and role-playing. Teachers also are provided with a manual describing the program.

Cost

The manual for each grade-level reading version of *Peer-Assisted Learning Strategies* costs from \$40 to \$44 (rates effective 06/13/2011). It includes teaching scripts and master copies of student materials. Video (or DVD) materials that provide an overview of the kindergarten and grades 2–6 programs are available for \$15 each. For an onsite one-day teacher-training workshop, the presenter's fee is estimated at \$1,500 plus travel expenses. Additional information can be found on the *Peer-Assisted Learning Strategies* website (<http://kc.vanderbilt.edu/pals>). Cost information for *Peer-Assisted Literacy Strategies* is available from the developer.

Research Summary

Forty-five studies reviewed by the WWC investigated the effects of *PALS* on beginning readers. Two studies (McMaster, Fuchs, Fuchs, & Compton, 2005; Stein, Berends, Fuchs, McMaster, Sáenz, Yen, & Compton, 2008) are randomized controlled trials that meet WWC evidence standards. One study (Mathes & Babyak, 2001) is a randomized controlled trial that meets WWC evidence standards with reservations. These three studies are summarized in this report. The remaining 42 studies do not meet either WWC eligibility screens or evidence standards. (See references beginning on p. 7 for citations for all 45 studies.)

Two additional studies were reviewed against the pilot Single-Case Design standards. Both studies do not meet pilot Single-Case Design standards. Studies reviewed against pilot Single-Case Design standards are listed in Appendix D and do not contribute to the intervention's rating of effectiveness.

Table 2. Scope of reviewed research

Grade	K, 1
Delivery method	Small group
Program type	Supplement
Studies reviewed	45
Meets WWC standards	2 studies
Meets WWC standards with reservations	1 study

Summary of studies meeting WWC evidence standards without reservations

McMaster et al. (2005) examined the effects of *Peer-Assisted Learning Strategies* and modified *Peer-Assisted Learning Strategies*⁷ compared to one-on-one adult tutoring. The participants in this study were 66 first graders in eight schools in Nashville, Tennessee who had been exposed to *Peer-Assisted Learning Strategies* for seven weeks and who scored 0.5 standard deviations or more below average readers on the curriculum-based measurement measures. These students were randomly assigned to one of the two *Peer-Assisted Learning Strategies* conditions or to adult tutoring. The WWC based its effectiveness rating on findings from 41 first-grade students in the *Peer-Assisted Learning Strategies* and adult-tutoring conditions.⁸ The study reported student outcomes after 13 weeks of program implementation.

Stein et al. (2008) included 2,959 kindergarten students in 67 schools in Nashville, Tennessee; Minnesota; and south Texas. Teachers were randomly assigned either to one of three *Peer-Assisted Learning Strategies* conditions that differed by the amount of training and support offered to the teacher or to a business-as-usual control group. These conditions included a day-long training workshop, the workshop plus two follow-up booster sessions, or the workshop and booster sessions plus weekly technical assistance provided by a graduate student. The study reported student outcomes after approximately 20 weeks of program implementation.

Summary of study meeting WWC evidence standards with reservations

Mathes and Babyak (2001) included 130 first-grade students from five schools in a medium-sized school district in Florida. The researchers compared two interventions (*Peer-Assisted Literacy Strategies* and *Peer-Assisted Literacy Strategies plus Mini-Lessons*) to a comparison group that used a typical reading curriculum with no supplement. Teachers were matched on demographic characteristics to form a stratified sample and then randomly assigned either to one of the two intervention groups or to the comparison group. After random assignment, teachers selected five students with different achievement levels within each participating classroom to be part of the study. Because the selection of students by teachers was not random, the analysis cannot be considered a randomized controlled trial. However, since the groups were shown to be alike on the comprehension measure at the start of the intervention, the study meets evidence standards with reservations.⁹ The study reported student outcomes after 14 weeks of program implementation.

Effectiveness Summary

The WWC review of interventions for Beginning Reading addresses student outcomes in four domains: alphabets, fluency, comprehension, and general reading achievement. The three studies that contribute to the effectiveness ratings in this report cover three domains: alphabets, fluency, and comprehension. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *PALS* on beginning readers. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 22.

Summary of effectiveness for the alphabets domain

Two studies reported findings in the alphabets domain.

McMaster et al. (2005) reported no statistically significant difference between the *PALS* group and the comparison group of first graders on the seven alphabets outcomes: Blending, Rapid Letter Naming, Rapid Letter Sound, Word Attack, Word Identification, Spelling, and Segmentation. The WWC-calculated effects were not statistically significant, and the average effect across the seven measures was not large enough to be considered substantively important according to WWC criteria (i.e., an effect size of at least 0.25).

Stein et al. (2008) reported, and the WWC confirmed, a statistically significant positive difference between the *K-PALS* group and the comparison group for each of the three *K-PALS* conditions (workshop, workshop plus booster, and workshop plus booster plus helper) on Rapid Letter Sounds for kindergarten students.

Thus, for the alphabets domain, one study shows statistically significant positive effects and one study shows indeterminate effects. This results in a rating of potentially positive effects, with a medium to large extent of evidence.

Table 3.1. Rating of effectiveness and extent of evidence for the alphabets domain

Rating of effectiveness	Criteria met
Potentially positive effects <i>Evidence of a positive effect with no overriding contrary evidence.</i>	The review of <i>PALS</i> in the alphabets domain had one study showing statistically significant positive effects, one study showing indeterminate effects, and no studies showing a statistically significant or substantively important negative effect.
Extent of evidence	Criteria met
Medium to large	The review of <i>PALS</i> in the alphabets domain was based on two studies that included 75 schools and 3,000 students.

Summary of effectiveness for the fluency domain

One study reported findings in the fluency domain.

McMaster et al. (2005) did not find statistically significant effects on the Near-Transfer Fluency and Far-Transfer Fluency outcomes for students in grade 1. The WWC-calculated effects were not statistically significant, and the average effect across the two measures was not large enough to be considered substantively important according to WWC criteria.

Thus, for the fluency domain, one study shows indeterminate effects. This results in a rating of no discernible effects, with a small extent of evidence.

Table 3.2. Rating of effectiveness and extent of evidence for the fluency domain

Rating of effectiveness	Criteria met
No discernible effects <i>No affirmative evidence of effects.</i>	The review of <i>PALS</i> in the fluency domain had one study showing indeterminate effects.
Extent of evidence	Criteria met
Small	The review of <i>PALS</i> in the fluency domain was based on one study that included eight schools and 41 students.

Summary of effectiveness for the comprehension domain

Two studies reported findings in the comprehension domain.

McMaster et al. (2005) reported a negative mean difference between the *PALS* group and the comparison group on the Comprehension subtest of the Comprehensive Reading Assessment Battery (CRAB) for students in grade 1. The WWC-calculated effect was not statistically significant, but it was large enough to be considered substantively important according to WWC criteria.

For Mathes and Babyak (2001), the WWC computed differences on the Woodcock Reading Mastery Tests–Revised (WRMT-R) Passage Comprehension subtest for the combined *PALS* and *PALS plus Mini-Lessons* conditions and the comparison condition for students in grade 1. The WWC-calculated effect was not statistically significant, but it was large enough to be a substantively important positive effect (i.e., an effect size of at least 0.25).

Thus, for the comprehension domain, one study found a substantively important negative effect, and one study found a substantively important positive effect. This results in a rating of mixed effects, with a small extent of evidence.

Table 3.3. Rating of effectiveness and extent of evidence for the comprehension domain

Rating of effectiveness	Criteria met
Mixed effects <i>Evidence of inconsistent effects.</i>	The review of <i>PALS</i> in the comprehension domain had one study showing substantively important negative effects and one study showing substantively important positive effects.
Extent of evidence	Criteria met
Small	The review of <i>PALS</i> in the comprehension domain was based on two studies that included 13 schools and 171 students.

References

Studies that meet WWC evidence standards without reservations

- McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children, 71*(4), 445–463.
- Stein, M. L., Berends, M., Fuchs, D., McMaster, K., Sáenz, L., Yen, L., & Compton, D. L. (2008). Scaling up an early reading program: Relationships among teacher support, fidelity of implementation, and student performance across different sites and years. *Educational Evaluation and Policy Analysis, 30*(4), 368–388.

Study that meets WWC evidence standards with reservations

- Mathes, P. G., & Babyak, A. E. (2001). The effects of Peer-Assisted Literacy Strategies for first-grade readers with and without additional mini-skills lessons. *Learning Disabilities Research & Practice, 16*(1), 28–44.

Studies that do not meet WWC evidence standards

- Baker, R. H. (2005). Teacher-directed instruction plus classwide peer tutoring and the reading growth of first-grade students. *Masters Abstracts International, 44*(02), 59-1053. (AAI1428630) 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.
- Bergeron, J. (1998). A comparison of classwide cross-age and same-age peer tutoring for second-grade students at risk for reading failure. *Dissertation Abstracts International, 59*(09), 3390A. (UMI No. 9905010) 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.
- Calhoon, M. B., Al Otaiba, S., Cihak, D., King, A., & Avalos, A. C. (2007). Effects of a peer-mediated program on reading skill acquisition for two-way bilingual first grade classrooms. *Learning Disability Quarterly, 30*(3), 169–184. The study does not meet WWC evidence standards because it is a randomized controlled trial in which the combination of overall and differential attrition rates exceeds WWC standards for this area, and the subsequent analytic intervention and comparison groups are not shown to be equivalent.
- Calhoon, M. B., Al Otaiba, S., Greenberg, D., King, A., & Avalos, A. (2007). Improving reading skills in predominantly Hispanic Title 1 first-grade classrooms: The promise of Peer-Assisted Learning Strategies. *Learning Disabilities Research & Practice, 21*(4), 261–272. The study does not meet WWC evidence standards because it is a randomized controlled trial in which the combination of overall and differential attrition rates exceeds WWC standards for this area, and the subsequent analytic intervention and comparison groups are not shown to be equivalent.
- Fuchs, L. S., Fuchs, D., Kazdan, S., & Allen, S. (1999). Effects of peer-assisted learning strategies in reading with and without training in elaborated help giving. *The Elementary School Journal, 99*(3), 201–219. 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.
- Mathes, P. G., Howard, J. K., Allen, S. H., & Fuchs, D. (1998). Peer-assisted learning strategies for first-grade readers: Responding to the needs of diverse learners. *Reading Research Quarterly, 33*(1), 62–94. The study does not meet WWC evidence standards because the estimates of effects did not account for differences in pre-intervention characteristics while using a quasi-experimental design.
- Mathes, P. G., Torgesen, J. K., Clancy-Menchetti, J., Santi, K., Nicholas, K., Robinson, C., & Grek, M. (2003). A comparison of teacher-directed versus peer-assisted instruction to struggling first-grade readers. *The Elementary School Journal, 103*(5), 459–479. 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.

Wehby, J. H., Falk, K. B., Barton-Arwood, S., Lane, K. L., & Cooley, C. (2003). The impact of comprehensive reading instruction on the academic and social behavior of students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders*, 11(4), 225. 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.

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

Barton-Arwood, S. M., Wehby, J. M., & Falk, K. B. (2005). Reading instruction for elementary age students with emotional and behavioral disorders: Academic and behavioral outcomes. *Exceptional Children*, 72(1), 7–27. 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.

Deshler, D. D., Palincsar, A. S., Biancarosa, G., & Nair, M. (2007). *Informed choices for struggling adolescent readers: A research-based guide to instructional programs and practices*. Newark, DE: International Reading Association. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Dion, E., Fuchs, D., & Fuchs, L. S. (2005). Differential effects of Peer-Assisted Learning Strategies on students' social preference and friendship making. *Behavioral Disorders*, 30(4), 421–429. The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.

Dunn, B. N. (2009). *PALS: Peer Assisted Learning Strategies* (ED507452). Education Resources Information Center. Online submission. Retrieved from <http://www.eric.ed.gov/PDFS/ED507452.pdf>. 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.

Eckert, T. L., Coddling, R. M., Truckenmiller, A. J., & Rheinheimer, J. L. (2009). Improving children's fluency in reading, mathematics, spelling, and writing: A review of evidence-based academic interventions. In A. Akin-Little, S. G. Little, M. A. Bray, & T. J. Kehle (Eds.), *Behavioral interventions in schools: Evidence-based positive strategies* (pp. 111–124). Washington, DC: American Psychological Association. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, 79(1), 262–300. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Falk, K. B., & Wehby, J. H. (2001). The effects of peer-assisted learning strategies on the beginning reading skills of young children with emotional or behavioral disorders. *Behavioral Disorders*, 26(4), 344–359. 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.

Foorman, B., & Al Otaiba, S. (2009). Reading remediation: State of the art. In K. Pugh & P. McCardle (Eds.), *How children learn to read: Current issues and new directions in the integration of cognition, neurobiology and genetics of reading and dyslexia research and practice* (pp. 257–274). New York: Psychology Press. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Fuchs, D., & Fuchs, L. S. (2007). Increasing strategic reading comprehension with peer-assisted learning activities. In D. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 175–199). New York: Psychology Press. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

- Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. (1997). Peer-Assisted Learning Strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34(1), 174–206. 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.
- Fuchs, L. S., Fuchs, D., & Kazdan, S. (1999). Effects of peer-assisted learning strategies on high school students with serious reading problems. *Remedial and Special Education*, 20(5), 309–318. 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.
- Ginsburg-Block, M., Rohrbeck, C., Lavigne, N., & Fantuzzo, J. W. (2008). Peer-assisted learning: An academic strategy for enhancing motivation among diverse students. In C. Hudley & A. E. Gottfried (Eds.), *Academic motivation and the culture of school in childhood and adolescence* (pp. 247–273). New York: Oxford University Press. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Han, I. (2010). Evidence-based reading instruction for English language learners in preschool through sixth grades: A meta-analysis of group design studies. *Dissertation Abstracts International*, 70(9A), 3412. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Hudson, K. G. (2004). The effects of Peer-Assisted Learning Strategies on the reading achievement of elementary students with and without decoding weaknesses. *Dissertation Abstracts International*, 65(10), 3754A. (UMI No. 3149163) 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.
- Joseph, L. M., & Schisler, R. (2009). Should adolescents go back to the basics? A review of teaching word reading skills to middle and high school students. *Remedial and Special Education*, 30(3), 131–147. 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.
- Keaton, J. M., Palmer, B. C., Nicholas, K. R., & Lake, V. E. (2007). Direct instruction with playful skill extensions: Action research in emergent literacy development. *Reading Horizons*, 47(3), 229–250. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Kroeger, S. D., Burton, C., & Preston, C. (2009). Integrating evidence-based practices in middle science reading. *Teaching Exceptional Children*, 41(3), 6–15. 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.
- Largent, J. (2009). *Response to intervention: Effectiveness of reducing the gap of reading scores of students enrolled in a Peer Assisted Learning Strategy (PALS) in fifth grade in a selected Missouri school district* (Unpublished research paper). Northwest Missouri State University, Maryville. 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.
- Mackley, S. R. (2009). *Kindergarten Peer-Assisted Learning Strategies with English language learners: An empirical dissertation* (Unpublished doctoral thesis). Philadelphia College of Osteopathic Medicine, 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.
- Mather, N., & Urso, A. (2008). Teaching younger readers with reading difficulties. In R. J. Morris & N. Mather (Eds.), *Evidence-based interventions for students with learning and behavioral challenges* (pp. 163–192). New York: Taylor & Francis Group. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Mathes, P. G., Torgesen, J. K., & Allor, J. H. (2001). The effects of Peer-Assisted Literacy Strategies for first-grade readers with and without additional computer-assisted instruction in phonological awareness. *American Educational Research Journal*, 38(2), 371–410. The study is ineligible for review because it does not use a comparison group design or a single-case design.

- McFay, D. D. (2008). *Response-to-intervention: Use of curriculum-based measurements to facilitate student achievement for learners at-risk of academic failure in reading at the middle school level* (Unpublished doctoral dissertation). University of Georgia, Athens. 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.
- McMaster, K. L., Kung, H., Han, I., & Cao, M. (2008). Peer-Assisted Learning Strategies: A “tier 1” approach to promoting English learners’ response to intervention. *Exceptional Children, 74*(2), 194–214. 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.
- Pearson, J. M. (2004). The effect of peer-assisted literacy strategies on the social standing of first-grade readers. *Dissertation Abstracts International, 65*(02), 412A. (UMI No. 3122359) The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.
- Petursdottir, A., McMaster, K., McComas, J. J., Bradfield, T., Braganza, V., Koch-McDonald, J.,...Scharf, H. (2009). Brief experimental analysis of early reading interventions. *Journal of School Psychology, 47*(4), 215–243. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Rathvon, N. (2008). *Effective school interventions: Evidence-based strategies for improving student outcomes* (2nd ed.). New York: Guilford Press. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Reed, D., & Vaughn, S. (2010). Reading interventions for older students. In T. A. Glover & S. Vaughn (Eds.), *The promise of response to intervention: Evaluating current science and practice* (pp. 143–186). New York: Guilford Press. 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.
- Rutherford, L. E., DuPaul, G. J., & Jitendra, A. K. (2008). Examining the relationship between treatment outcomes for academic achievement and social skills in school-age children with attention-deficit hyperactivity disorder. *Psychology in the Schools, 45*(2), 145–157. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Sáenz, L. M., Fuchs, L. S., & Fuchs, D. (2005). Peer-Assisted Learning Strategies for English language learners with learning disabilities. *Exceptional Children, 71*(3), 231–247. 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., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis. *Reading Research Quarterly, 43*(3), 290–322. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). *Effective beginning reading programs: A best-evidence synthesis*. Baltimore, MD: Johns Hopkins University, Center for Data-Driven Reform in Education. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research, 79*(4), 1391–1466. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Spörer, N., & Brunstein, J. C. (2009). Fostering the reading comprehension of secondary school students through peer-assisted learning: Effects on strategy knowledge, strategy use, and task performance. *Contemporary Educational Psychology, 34*(4), 289–297. 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.
- Wayman, M. M., McMaster, K., Sáenz, L., & Watson, J. (2010). Using curriculum-based measurement to monitor secondary English language learners’ responsiveness to peer-mediated reading instruction. *Reading & Writing Quarterly, 26*(4), 308–332. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Appendix A.1: Research details for McMaster et al. (2005)

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children, 71*(4), 445–463.

Table A1. Summary of findings

Meets WWC evidence standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index (percentile points)	Statistically significant
Alphabetics	8 schools/41 students	–6	No
Fluency	8 schools/41 students	–8	No
Comprehension	8 schools/41 students	–12	No

Setting Eight elementary schools in metropolitan Nashville, Tennessee, participated in the study. Four of the eight study schools were classified as Title I schools; the other four were middle-class, non-Title I schools.

Study sample Before the study began, 323 first-grade students used *PALS* for seven weeks and were subsequently tested. The 66 students who scored 0.5 standard deviations or more below average readers in terms of both level and slope on the curriculum-based measures comprised the sample for this study. These 66 students were randomly assigned to one of three conditions: *PALS*, modified *PALS*, and adult tutoring, with 22 in each condition.

The intervention was delivered one-on-one by peers in the *PALS* and modified *PALS* conditions and one-on-one by up to eight adult tutors in the adult-tutoring condition. The final analysis sample consisted of 56 students: 21 in *PALS*, 15 in modified *PALS*, and 20 in tutoring. Fifteen (27%) of the 56 students were English language learner students. Only the *PALS* vs. tutoring analysis meets WWC evidence standards, so the analysis sample used in this review includes 41 students in eight schools.

Intervention group *PALS* is a peer-tutoring program that emphasizes phonological awareness, decoding, and fluency. In this study, it was implemented three times a week for 35 minutes each session. Teachers paired higher performing and lower performing readers who took turns coaching each other. The intervention group received *PALS* over the course of 13 weeks.

Comparison group The comparison group received one-on-one tutoring from trained adult research assistants. Adult tutoring took place three times a week, 35 minutes each session, for 13 weeks, and covered the same topics as in the two *PALS* conditions. The tutoring session was structured similar to a special education pullout program, with greater attention to skill mastery and the student’s specific needs. The study viewed *PALS* as the business-as-usual comparison group, but the WWC treated the tutoring condition as the comparison for the purposes of this review.

Outcomes and measurement

Testing was conducted at baseline and at follow-up by two full-time project coordinators and eight graduate students who were trained to ensure inter-rater agreement of at least 90%. Students were tested over two one-on-one sessions in a quiet location in their school. Students were not tested by staff who had tutored them. The baseline Dolch measure, developed by research staff, was used as the covariate in analyses. For this measure, the score was recorded as the number of high-frequency words read correctly in one minute. The outcomes included in this study were Blending, Rapid Letter Naming, Rapid Letter Sound, Segmentation, Spelling, Word Identification, and Word Attack in the alphabetic domain; Near-Transfer Fluency and Far-Transfer Fluency in the fluency domain; and the Comprehensive Reading Assessment Battery: Comprehension measure in the comprehension domain. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Teachers were trained to use *PALS* in October through a one-day training session before the start of this study. A research staff member visited each classroom twice weekly over the seven-week period of the initial *PALS* implementation. In January, research staff attended a one-day workshop to learn the modified *PALS* and tutoring procedures. Each staff member was then assigned to implement the tutoring or modified *PALS* intervention. Intervention fidelity for *PALS* was measured at 92% based on classroom checks conducted in December and March.

Appendix A.2: Research details for Stein et al. (2008)

Stein, M. L., Berends, M., Fuchs, D., McMaster, K., Sáenz, L., Yen, L., & Compton, D. L. (2008). Scaling up an early reading program: Relationships among teacher support, fidelity of implementation, and student performance across different sites and years. *Educational Evaluation and Policy Analysis, 30*(4), 368–388.

Table A2. Summary of findings

Meets WWC evidence standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index (percentile points)	Statistically significant
Alphabetic	67 schools/2,959 students	+30	Yes

Setting

The study took place in 71 schools in three sites over two years: 14 schools in Nashville, Tennessee; 36 schools in Minnesota; and 21 schools in south Texas. The final analytic sample included 67 schools.

Study sample

Project staff first recruited schools to obtain balanced samples on site-specific factors: Title I status in Nashville, Title I status and whether the school offered half-day or full-day kindergarten in Minnesota, and the proportion of limited English proficiency students in the schools in south Texas. Teachers were recruited within the selected schools, and 224 teachers participated over the two study years (55 teachers participated in both years, for a total of 279 teacher-years). Within each participating school, teachers were randomly assigned to one of the four conditions: control, workshop, workshop plus booster, or workshop plus booster and helper. The study does not report the number of teachers in each condition.

Study sample (continued)

Researchers obtained parental consent for more than 90% of the students in the classrooms of study teachers. These students were pretested, and 12 students were selected from each class: four children with the lowest reading scores, four children with the highest scores, and four children with scores in the middle of the score distribution. The consented study sample included 3,171 kindergarten students, with 668 in the control condition, 968 in the workshop condition, 931 in the booster condition, and 604 in the helper condition. The final hierarchical linear modeling (HLM) analysis sample included only 2,959 students and 259 teachers. The WWC could not calculate attrition by condition based on the information provided in the study. However, based on reasonable assumptions about how to attribute overall attrition to groups, the study is assumed to have low differential attrition.

Twenty-four percent of the students in the study were English language learners, 62% were eligible for free or reduced-price lunch, 50% were female, and 5% had Individualized Education Plans. Of the students in the sample, 40% were Hispanic, 26% were non-Hispanic White, 25% were African American, 5% were Asian, and 3% were of other ethnicities.

Intervention group

The study included three treatment conditions: (1) a day-long training workshop (*K-PALS*), (2) the workshop plus two follow-up booster sessions (*K-PALS + Booster*), and (3) the workshop and booster sessions plus weekly technical assistance provided by a graduate student (*K-PALS + Booster + Helper*). Although the treatment conditions vary by the amount of training and support received by teachers, the *K-PALS* intervention was the same in all three treatment conditions. Students were paired by their teachers and then worked through structured lessons during 35-minute sessions implemented four times per week in this study. Stronger readers were paired with weaker readers, and pairings were maintained for four to six weeks before being reorganized. Within each pair, students took turns acting as the reader and the coach. The classroom teacher monitored the pairs and provided feedback as necessary. Program materials, including a teacher manual and all student worksheets, were provided by *K-PALS*.

Comparison group

The comparison was a business-as-usual counterfactual. Comparison teachers did not implement the intervention and did not receive any additional training.

Outcomes and measurement

The primary outcome is Rapid Letter Sounds, an alphabetic measure developed by Levy and Lysunchuk (1997).¹⁰ All study students were tested approximately three weeks before the intervention began and again 20 weeks later. For a more detailed description of the outcome measure, see Appendix B.

Support for implementation

All teachers in the three treatment groups attended a day-long training workshop before the intervention began. For the *K-PALS + Booster* treatment group, two follow-up booster sessions were also provided to allow teachers to review program procedures and to identify and solve implementation issues. Teachers in the *K-PALS + Booster + Helper* treatment group attended the training workshop and booster sessions and also had weekly technical assistance provided by a trained graduate assistant. Average implementation fidelity, measured at two points during implementation by the project coordinator, was 86%.

Appendix A.3: Research details for Mathes & Babyak (2001)

Mathes, P. G., & Babyak, A. E. (2001). The effects of Peer-Assisted Literacy Strategies for first-grade readers with and without additional mini-skills lessons. *Learning Disabilities Research & Practice*, 16(1), 28–44.

Table A3. Summary of findings

Meets WWC evidence standards with reservations

Outcome domain	Sample size	Study findings	
		Average improvement index (percentile points)	Statistically significant
Comprehension	5 schools/130 students	+17	No

Setting The study took place in five schools in a medium-sized school district in Florida.

Study sample Thirty first-grade teachers from five schools matched on demographic characteristics were selected to form a stratified sample and were randomly assigned to one of three conditions: *Peer-Assisted Literacy Strategies (PALS; 10 teachers)*, *Peer-Assisted Literacy Strategies plus Mini-Skills Lesson (PALS plus ML; 10 teachers)*, or a comparison group (10 teachers). After rank-ordering students by their reading ability within the classroom, each teacher identified five students to be included in the analysis sample: one high-achieving student, one average-achieving student, and three low-achieving students. High- and average-achieving students from the *PALS plus ML* group did not participate in the *ML* component of the intervention; thus, sample sizes for the *PALS plus ML* group are smaller than the other groups. The study began with 150 first-grade students. After attrition, the final analysis sample was 130 students (61 students in *PALS*, 20 in *PALS plus ML*, and 49 in the comparison group) and 28 teachers.

The post-attrition samples were checked for equivalence at pretest by the WWC, and only one comparison was found to be comparable. The WWC intervention rating is based on the comparison of the combined *PALS* and *PALS plus ML* conditions to the comparison group with a total of 130 students across different ability groups. The mean age of the participating students was 6.9 years. Forty-seven percent of the students were female, 39% were African American, 59% were White, and 32% had special needs.¹¹

Intervention group This study included two intervention conditions, *PALS* and *PALS plus ML*, that are combined for the purposes of this review. Teachers in both of these groups implemented *PALS* with their entire class for 14 weeks in 35-minute sessions three times a week. In each lesson, a stronger reader and a weaker reader were paired. In Sounds and Words activities, students practiced phonemic segmentation, applied alphabetic knowledge to decoding novel words, and read connected text built on previously mastered phonological elements. During Story Sharing time, students made predictions about a book prior to reading it, shared the experience of reading the book with a peer, had repeated exposure to the text, and summarized the text through verbal retelling. In the *PALS plus ML* condition, a 15- to 20-minute mini-lesson was also given to small groups of low-achieving students in each classroom three times a week during the last six weeks of the *PALS* intervention. Teachers taught the mini-lessons before the *PALS* sessions. The content of the mini-lessons was the same as the Words and Sounds portion of *PALS*.

Comparison group

Teachers used their regular reading curriculum and did not receive any recommendations or feedback about instruction from the researchers. However, *PALS* staff collected student data weekly using the Continuous Progress Monitoring measure across all groups (treatment and comparison). All teachers were also given a graph showing students' progress every month.

Outcomes and measurement

The study included several outcome measures, but only the analysis of achievement using the Woodcock Reading Mastery Test–Revised Passage Comprehension subtest meets WWC evidence standards with reservations. For a more detailed description of this outcome measure, see Appendix B.

Support for implementation

Intervention teachers participated in an all-day in-service workshop prior to the intervention. They were provided with a manual describing *PALS* and practiced using the intervention. During training, *PALS* project staff were available to provide support needed to implement the program. Project staff conducted three observations of teachers and students; intervention fidelity was 93% for teachers, and ranged from 75% (Sounds and Words) to 82% (Story Sharing) for students.

Appendix B: Outcome measures for each domain

Alphabetics	
Phonemic awareness construct	
<i>Blending</i>	This blending task, developed by Fuchs et al. (2001), awards one point for each word blended correctly from phonemes read by the examiner (as cited in McMaster et al., 2005).
<i>Segmentation</i>	This segmentation task, developed by Fuchs et al. (2001), measures the number of phonemes expressed correctly in one minute. This test is based on the Yopp-Singer test (as cited in McMaster et al., 2005).
Letter naming construct	
<i>Rapid Letter Naming</i>	This measure, developed by Fuchs et al. (2001), records the number of letters named correctly in one minute by the student from a set of uppercase and lowercase letters typed randomly on a sheet of paper (as cited in McMaster et al., 2005).
Phonics construct	
<i>Rapid Letter Sound</i>	This test measures the number of letter sounds pronounced correctly in one minute by the student from a sheet that shows the 26 letters of the alphabet in random order (as cited in McMaster et al., 2005; Stein et al., 2008).
<i>Wechsler Individual Achievement Test: Spelling subtest</i>	This standardized test measures the number of words and letters written correctly by the student (as cited in McMaster et al., 2005).
<i>Woodcock Reading Mastery Tests—Revised (WRMT-R): Word Attack subtest</i>	The Word Attack subtest is a measure of phonemic reading ability in which the student reads nonwords. This is a standardized test with 51 items (as cited in McMaster et al., 2005).
<i>WRMT-R: Word Identification subtest</i>	The Word Identification subtest is a measure of word reading vocabulary in which the student reads a list of words of increasing difficulty. This is a standardized test (as cited in McMaster et al., 2005).
Fluency	
<i>Near-Transfer Fluency</i>	This test measures the number of words read correctly in one minute from two passages that were similar to the passages used in the <i>PALS</i> lessons (as cited in McMaster et al., 2005).
<i>Far-Transfer Fluency</i>	This test measures the number of words read correctly in one minute. The two passages in this test were traditional folktales selected from the Comprehensive Reading Assessment Battery (as cited in McMaster et al., 2005).
Comprehension	
Reading comprehension construct	
<i>Comprehensive Reading Assessment Battery: Comprehension</i>	This test measures comprehension of the two passages read for the Far-Transfer Fluency subtest using 10 open-ended questions (as cited in McMaster et al., 2005).
<i>WRMT-R: Passage Comprehension subtest</i>	In this standardized test, comprehension is measured by having students read silently and fill in missing words in a short paragraph (as cited in Mathes & Babyak, 2001).

Appendix C.1: Findings included in the rating for the alphabetics domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
McMaster et al., 2005^a								
<i>Blending</i>	Grade 1	8 schools/ 41 students	19.30 (8.54)	22.77 (7.14)	-3.47	-0.43	-17	> 0.05
<i>Segmentation</i>	Grade 1	8 schools/ 41 students	35.59 (13.39)	35.19 (14.04)	0.40	0.03	+1	> 0.05
<i>Rapid Letter Naming</i>	Grade 1	8 schools/ 41 students	48.24 (17.11)	45.03 (19.19)	3.21	0.17	+7	> 0.05
<i>Rapid Letter Sound</i>	Grade 1	8 schools/ 41 students	42.29 (11.15)	44.95 (15.54)	-2.66	-0.19	-8	> 0.05
<i>WIAT: Spelling</i>	Grade 1	8 schools/ 41 students	12.76 (3.37)	12.45 (2.86)	0.31	0.10	+4	> 0.05
<i>WRMT-R: Word Attack</i>	Grade 1	8 schools/ 41 students	6.88 (5.26)	8.79 (5.36)	-1.91	-0.35	-14	> 0.05
<i>WRMT-R: Word Identification</i>	Grade 1	8 schools/ 41 students	21.15 (9.56)	25.09 (9.51)	-3.94	-0.41	-16	> 0.05
Domain average for alphabetics (McMaster et al., 2005)						-0.15	-6	Not statistically significant
Stein et al., 2008^b								
Comparison #1: K-PALS vs. control								
<i>Rapid Letter Sound</i>	Kindergarten	67 schools/ 1,636 students	44.00 (18.90)	32.90 (17.00)	11.10	0.61	+23	< 0.01
Comparison #2: K-PALS + Booster vs. control								
<i>Rapid Letter Sound</i>	Kindergarten	67 schools/ 1,599 students	51.57 (17.80)	32.90 (17.00)	18.67	1.07	+36	< 0.01
Comparison #3: K-PALS + Booster + Helper vs. control								
<i>Rapid Letter Sound</i>	Kindergarten	67 schools/ 1,272 students	49.01 (19.10)	32.90 (17.00)	16.11	0.89	+31	< 0.01
Domain average for alphabetics (Stein et al., 2008)						0.86	+30	Statistically significant
Domain average for alphabetics across all studies						0.35	+14	na

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study's domain average was determined by the WWC. Stein et al. (2008) is characterized as having a statistically significant positive effect when univariate statistical tests are reported for each outcome measure, the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant. WIAT = Wechsler Individual Achievement Test. WRMT-R = Woodcock Reading Mastery Tests-Revised. na = not applicable.

^a For McMaster et al. (2005), a correction for multiple comparisons was needed but did not affect significance levels. The *p*-values presented here were reported in the original study. The group means presented here are ANCOVA-adjusted posttest measures for which a pre-intervention Dolch measure was used as the covariate. The group means for spelling are calculated using a difference-in-differences approach (see the WWC Procedures and Standards Handbook, Appendix B); the WWC calculated the program group mean 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 mean.

^b For Stein et al. (2008), a correction for multiple comparisons was needed but did not affect significance levels. The *p*-values presented here were reported in the original study. The study does not provide the exact breakdown of students in the final HLM analytic sample across conditions (*n* = 2,959), so the number of students (*n* = 3,171) listed here is from Table 1 (p. 376). The control group means shown are unadjusted posttest; the intervention group means shown are the sum of the unadjusted control group means and the HLM level-2 coefficient reported in the study.

Appendix C.2: Findings included in the rating for the fluency domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			<i>p</i> -value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
McMaster et al., 2005^a								
<i>Near-Transfer Fluency</i>	Grade 1	8 schools/ 41 students	18.95 (12.55)	21.54 (10.68)	-2.59	-0.22	-9	> 0.05
<i>Far-Transfer Fluency</i>	Grade 1	8 schools/ 41 students	20.01 (12.22)	22.27 (10.08)	-2.26	-0.20	-8	> 0.05
Domain average for fluency (McMaster et al., 2005)						-0.21	-8	Not statistically significant

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study's domain average was determined by the WWC.

^a For McMaster et al. (2005), a correction for multiple comparisons was needed but did not affect significance levels. The *p*-values presented here were reported in the original study. The group means presented here are ANCOVA-adjusted posttest measures for which a pre-intervention Dolch measure was used as the covariate.

Appendix C.3: Findings included in the rating for the comprehension domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			<i>p</i> -value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
McMaster et al., 2005^a								
<i>CRAB: Comprehension</i>	Grade 1	8 schools/ 41 students	0.42 (0.55)	0.64 (0.82)	-0.22	-0.31	-12	> 0.05
Domain average for comprehension (McMaster et al., 2005)						-0.31	-12	Not statistically significant
Mathes & Babyak, 2001^b								
<i>WRMT-R: Passage Comprehension</i>	Grade 1	5 schools/ 130 students	8.26 (7.30)	5.31 (5.50)	2.95	0.44	+17	0.07
Domain average for comprehension (Mathes & Babyak, 2001)						0.44	+17	Not statistically significant
Domain average for comprehension across all studies						0.06	+32	na

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study's domain average was determined by the WWC. CRAB = Comprehensive Reading Assessment Battery. WRMT-R = Woodcock Reading Mastery Test–Revised. na = not applicable.

^a For McMaster et al. (2005), no corrections for multiple comparisons or clustering were needed. The p -value presented here was reported in the original study. The group means presented here are ANCOVA-adjusted posttest measures for which a pre-intervention Dolch measure was used as the covariate. The original study included a second treatment condition, modified *PALS*. However, that condition had high attrition, and the remaining sample did not meet baseline equivalence and, therefore, does not meet evidence standards.

^b For Mathes and Babyak (2001), a correction for clustering was needed. The p -value presented here was computed by the WWC. The original study presented numerous comparisons defined by baseline achievement and condition (*PALS*, *PALS plus ML*, or comparison). The WWC checked baseline equivalence between these groups, as well as the equivalence of the groups pooled across all achievement levels. Of these, only the comparison of the control condition and the combined *PALS* conditions pooled across all student achievement levels meets standards with reservations. The numbers used in this evidence review refer to this pooled sample. The original study did not report the p -value for the pooled group. The WWC calculated the program group mean using a difference-in-differences approach (see the WWC Procedures and Standards Handbook, Appendix B) 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.

Appendix D: Single-case design studies reviewed for this intervention

Study citation	Study disposition
Lane, K. L., Little, M. A., Redding-Rhodes, J., Phillips, A., & Walsh, M. T. (2007). Outcomes of a teacher-led reading intervention for elementary students at risk for behavioral disorders. <i>Exceptional Children</i> , 74(1), 47–70.	Does not meet WWC pilot Single-Case Design standards because it does not have at least three attempts to demonstrate an intervention effect at three different points in time.
Lorah, K. S. (2003). Effects of peer tutoring on the reading performance and classroom behavior of students with attention deficit hyperactivity disorder. <i>Dissertation Abstracts International</i> , 64(04A), 198-1208.	Does not meet WWC pilot Single-Case Design standards because it does not have at least three attempts to demonstrate an intervention effect at three different points in time.

Table Notes: The supplemental studies presented in this table do not factor into the determination of the intervention rating.

Endnotes

¹ The descriptive information for this program was obtained from publicly available sources: the program's website (<http://kc.vanderbilt.edu/pals/>; downloaded September 2010); Fuchs, Fuchs, Kazdan, and Allen (1999); Mathes and Babyak (2001); Mathes, Howard, Allen, and Fuchs (1998); and Mathes et al. (2003). The WWC requests that developers review the program description sections for accuracy from their perspective. The program description was provided to the developer in June 2011; however the WWC received no response. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review. The literature search reflects documents publicly available by August 2011.

² *Peer-Assisted Learning Strategies* also includes program versions for high school and separate programs for mathematics that are outside the scope of this review.

³ This report has been updated to include reviews of 34 (group design) studies that have been reviewed since 2007. Of the additional studies, 28 were not within the scope of the review protocol, four were within the scope of the protocol but did not meet WWC evidence standards, and two met WWC evidence standards (McMaster et al., 2005, and Stein et al., 2008). The report confirmed the study disposition of meets standards with reservations for Mathes and Babyak (2001), the study included in the earlier report. Three studies that were included in the 2007 report—Fuchs et al. (1999), Mathes et al. (1998), and Mathes et al. (2003)—received revised dispositions in this report of do not meet WWC evidence standards. These revised dispositions are due to a change in the review protocol, particularly in baseline equivalence standards. A complete list and disposition of all studies reviewed are provided in the references. The studies in this report were reviewed using WWC Evidence Standards, version 2.1, as described in the Beginning Reading review protocol, version 2.0. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

⁴ For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 22. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies.

⁵ Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-Assisted Learning Strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34, 174–206.

⁶ *Peer-Assisted Literacy Strategies* also includes a teacher-directed version, not reviewed in this report, in which the teacher (or another adult) always serves as the tutor.

⁷ Compared to *PALS*, modified *PALS* incorporated some important alterations, such as placing greater emphasis on phonological awareness and decoding skill.

⁸ The modified *PALS* condition had high attrition; therefore, comparisons of this condition to the comparison group failed to meet WWC baseline equivalence standards.

⁹ Mathes and Babyak (2001) presented numerous comparisons defined by students' baseline achievement and treatment condition (*PALS*, *PALS plus ML*, or comparison). The WWC checked baseline equivalence between these groups, as well as the equivalence of the groups pooled across all achievement levels. Of these, only the comparison of the control condition and the combined *PALS* conditions pooled across all student achievement levels on the comprehension outcome meets WWC standards with reservations. The numbers used in this evidence review refer to this pooled sample.

¹⁰ Levy, B. A., & Lysunchuk, L. (1997). Beginning word recognition: Benefits of training by segmentation and whole word methods. *Scientific Studies of Reading*, 1, 359–387.

¹¹ This estimate corresponds to the row in the middle of Table 2 that starts with "referred to special education" (Mathes & Babyak, 2001, p. 32). The last row in the same section, "qualified," provides a more conservative estimate that corresponds to 44% of special education students in the analysis sample.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, May). *Beginning Reading intervention report: Peer-Assisted Learning/Literacy Strategies*. Retrieved from <http://whatworks.ed.gov>.

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC evidence standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC evidence standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
Positive effects	Two or more studies show statistically significant positive effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important negative effects.
Potentially positive effects	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
Mixed effects	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows a statistically significant or substantively important negative effect, but no more such studies than the number showing a statistically significant or substantively important positive effect, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
Potentially negative effects	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
Negative effects	Two or more studies show statistically significant negative effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important positive effects.
No discernible effects	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

Glossary of Terms

Attrition	Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.
Clustering adjustment	If treatment assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.
Confounding factor	A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.
Design	The design of a study is the method by which intervention and comparison groups were assigned.
Domain	A domain is a group of closely related outcomes.
Effect size	The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.
Eligibility	A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.
Equivalence	A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.
Extent of evidence	An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 22.
Improvement index	Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from -50 to +50.
Multiple comparison adjustment	When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.
Quasi-experimental design (QED)	A quasi-experimental design (QED) is a research design in which subjects are assigned to treatment and comparison groups through a process that is not random.
Randomized controlled trial (RCT)	A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into treatment and comparison groups.
Rating of effectiveness	The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 22.
Single-case design	A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.
Standard deviation	The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.
Statistical significance	Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).
Substantively important	A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the [WWC Procedures and Standards Handbook \(version 2.1\)](#) for additional details.