Appendix

Appendix A1 Study characteristics: Saenz, Fuchs, and Fuchs, 2005

Characteristic	Description
Study citation	Saenz, 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.
Participants	Twelve classrooms from grades 3–6 in one Texas school district were stratified based on grade level and school. Classrooms were then randomly assigned to either the *Peer-Assisted Learning Strategies* condition or the comparison condition. For a classroom to be eligible for the study, all students had to be English language learners, and at least two students had to have a learning disability (LD). Outcome data were collected on 11 students in each class: two students with LD, three low-achieving (LA) students, three average-achieving (AA) students, and three high-achieving (HA) students. The learning disability group is not included in this review since another WWC topic area will review those results. The students were categorized into LA, AA, and HA based on teachers' ranking according to classroom observations, previous scores on minimum state standards competency exams, and district-required informal reading inventories. LA students were in the lowest quartile of the class rank, AA in the middle half, and HA in the top quartile. The baseline sample included in this review consisted of 12 classrooms (six *Peer-Assisted Learning Strategies* and six comparison) and a total of 108 native Spanish-speaking students (54 *Peer-Assisted Learning Strategies* and 54 comparison) in grades 3–6. Of the 54 students in each condition, 18 were low achievers, 18 were average achievers, and 18 were high achievers. The analysis sample included in this review consisted of 12 classrooms (six *Peer-Assisted Learning Strategies* and six comparison) and 99 students (49 *Peer-Assisted Learning Strategies* and 50 comparison). Of the 49 *Peer-Assisted Learning Strategies* students in the analysis sample, 15 were low achievers, 17 were average achievers, and 18 were high achievers. Of the 50 comparison students in the analysis sample, 18 were low achievers, 18 were average achievers, and 14 were high achievers.
Setting	The study was conducted in one school district in Texas. All students were enrolled in bilingual education classrooms in grades 3–6.
Intervention	*Peer-Assisted Learning Strategies* sessions were conducted three times a week for 15 weeks. Each *Peer-Assisted Learning Strategies* session lasted for 25–35 minutes and occurred during regular reading instruction periods. Teachers ranked students by their reading achievement (high versus low) and paired a higher-achieving student with a lower-achieving student. Students were assigned a new partner about once a month. During *Peer-Assisted Learning Strategies*, pairs of students engaged in three reading activities: partner reading and retelling, paragraph shrinking, and prediction relay. In all three activities, students took 5-minute turns of being tutor and tutee. During partner reading and retelling, the better reader read aloud for five minutes while the weaker reader served as the tutor, who identified errors and corrected them. The weaker reader reread the same material for the next five minutes and retold what was read. During paragraph shrinking, each student read aloud for five minutes, stopping after each paragraph to summarize what was read. During prediction relay, the reader made a prediction before reading, read half a page, checked the prediction, and summarized using paragraph shrinking. Pairs earned points for correct or accurate responses during activities.[1]
Comparison	Teachers in the comparison group provided the district's regular curriculum for reading instruction. Lesson plans for both the intervention and comparison classrooms were reviewed twice during the study to assess the type of instruction provided. The study found that *Peer-Assisted Learning Strategies* teachers were more likely than comparison teachers to use one-on-one instruction, and no statistical differences were found in small-group instruction, whole-class instruction, and independent seatwork. The study found that *Peer-Assisted Learning Strategies* teachers were more likely than comparison teachers to use peer-mediated instruction and less likely to use teacher-led instruction.
Primary outcomes and measurement	The study measures in the reading achievement domain were three subtests of the Comprehensive Reading Assessment Battery. The subscales used were Word Correct, Maze Choices Correct, and Comprehension Questions Correct. For a more detailed description of these outcome measures, see Appendix A2.

1. Typically, the points are used as a motivation technique. Teachers can use them in several forms, for grades (participation points), prizes, class parties, and so on. In Saenz, Fuchs, and Fuchs (2005), the authors do not specify how these points were used by the teachers.

(continued)

Appendix A1 Study characteristics: Saenz, Fuchs, and Fuchs, 2005 *(continued)*

Characteristic	Description
Staff/teacher training	Teachers assigned to the *Peer-Assisted Learning Strategies* condition were trained by research assistants during a full-day workshop. Teachers were given an overview of *Peer-Assisted Learning Strategies* as well as opportunities to practice *Peer-Assisted Learning Strategies* procedures. Training emphasized how teachers could train their students to implement *Peer-Assisted Learning Strategies*. Upon conclusion of the workshop, teachers received a comprehensive *Peer-Assisted Learning Strategies* manual. The manual included scripted lessons that could be used when training students on *Peer-Assisted Learning Strategies* procedures. As part of this study, research assistants provided daily technical assistance to *Peer-Assisted Learning Strategies* teachers during the five weeks during which teachers trained students on *Peer-Assisted Learning Strategies* procedures. At the completion of student training, research assistants provided weekly technical assistance for the duration of *Peer-Assisted Learning Strategies* implementation.

Appendix A2 Outcome measures for the reading achievement domain[1]

Outcome measure	Description
Comprehensive Reading Assessment Battery (CRAB)[2]	The CRAB includes four 400-word folktales with a 2nd- to 3rd-grade readability level. Students have three minutes to read the first folktale aloud and then answer 10 comprehension questions. For a second folktale, students have two minutes to complete a cloze or maze task, three minutes to read the story aloud, and then answer 10 comprehension questions (as cited by Saenz, Fuchs, & Fuchs, 2005). CRAB has three subscales described below.
CRAB Words Correct subscale	This subscale assesses reading fluency and accuracy. Scores on this measure are based on the number of words read correctly in three minutes (as cited by Saenz, Fuchs, & Fuchs, 2005).
CRAB Maze Choices Correct subscale	The Maze Choices Correct subscale assesses silent reading accuracy and fluency. The maze task requires students to read a passage that consists of the first sentence intact, followed by every seventh word replaced with a three-item multiple-choice format. One answer is a semantically correct choice for the missing word. Scores on this measure are based on the number of correct maze choices made in two minutes (as cited by Saenz, Fuchs, & Fuchs, 2005).
CRAB Comprehension Questions Correct subscale	The Comprehension Questions Correct subscale assesses reading comprehension. Scores are based on the number of correct answers to comprehension questions (as cited by Saenz, Fuchs, & Fuchs, 2005).

1. This appendix reports outcome measures considered for the effectiveness rating and the average improvement indices for the reading achievement domain.
2. Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Monitoring reading growth using student recalls: Effects of two teacher feedback systems. *Journal of Educational Research, 83*, 103–111.

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

Outcome measure	Study sample	Sample size (clusters/ students)	Authors' findings from the study: Mean outcome[2] (standard deviation)[3] — *Peer-Assisted Learning Strategies* group[4]	Authors' findings from the study: Mean outcome[2] (standard deviation)[3] — Comparison group	WWC calculations: Mean difference[5] (*Peer-Assisted Learning Strategies* – comparison)	WWC calculations: Effect size[6]	WWC calculations: Statistical significance[7] (at $\alpha = 0.05$)	WWC calculations: Improvement index[8]
Saenz, Fuchs, and Fuchs, 2005[9]								
CRAB Words Correct subscale	Grades 3–6	12 classes/ 99 students	341.08 (82.49)	329.41 (88.37)	11.67	0.14	ns	+5
CRAB Comprehension Questions Correct subscale	Grades 3–6	12 classes/ 99 students	5.09 (2.39)	3.71 (1.61)	1.38	0.67	ns	+25
CRAB Maze Choices Correct subscale	Grades 3–6	12 classes/ 99 students	11.23 (4.21)	10.74 (3.81)	0.49	0.12	ns	+5
Average for reading achievement (Saenz, Fuchs, and Fuchs, 2005)[10]						**0.31**	**na**	**+12**

ns = not statistically significant

na = not applicable/not studied

CRAB = Comprehensive Reading Assessment Battery

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the reading achievement domain.
2. The means and standard deviations presented in this table for Saenz, Fuchs, and Fuchs (2005) were calculated by the WWC. Saenz, Fuchs, and Fuchs (2005) present separate results for low-achieving (LA), average-achieving (AA), and high-achieving (HA) subgroups at the class level. The results presented here are WWC aggregated results based on student-level findings presented in Saenz (2002) for the LA, AA, and HA subgroups.
3. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
4. The *Peer-Assisted Learning Strategies* group mean outcome values for Saenz, Fuchs, and Fuchs (2005) were calculated by the WWC using a difference-in-differences approach (see WWC 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.
5. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The mean difference for each of the reading achievement outcomes reported by Saenz, Fuchs, and Fuchs (2005) reflects the mean difference between treatment and control groups calculated by the WWC using the difference-in-differences approach.
6. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
7. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
8. 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.
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 Saenz, Fuchs, and Fuchs (2005), corrections for clustering and multiple comparisons were needed, so the significance levels may differ from those reported in the original study.
10. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.

Appendix A4 Summary of subgroup findings for the reading achievement domain[1]

Outcome measure	Study sample	Sample size (clusters/ students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *Peer-Assisted Learning Strategies* group[3]	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[4] (*Peer-Assisted Learning Strategies* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Saenz, Fuchs, and Fuchs, 2005[8]								
CRAB Words Correct subscale	Average achievers	12 classes/ 35 students	322.92 (66.98)	318.39 (75.32)	4.53	0.06	ns	+2
CRAB Comprehension Questions Correct subscale	Average achievers	12 classes/ 35 students	5.02 (1.76)	3.86 (1.37)	1.16	0.72	ns	+26
CRAB Maze Choices Correct subscale	Average achievers	12 classes/ 35 students	10.66 (3.66)	10.72 (3.18)	–0.06	–0.02	ns	–1

ns = not statistically significant
CRAB = Comprehensive Reading Assessment Battery

1. This appendix presents subgroup findings for measures that fall in the reading achievement domain. Total group scores are presented in Appendix A3. Saenz, Fuchs, and Fuchs (2005) present separate results for low-achieving (LA), average-achieving (AA), and high-achieving (HA) subgroups at the class level. The results presented here are based on student-level findings presented in Saenz (2002) for the AA subgroup. This report presents only results for the AA group since the LA and HA subgroup results failed to meet WWC evidence standards. The low-achieving subgroup results do not meet standards because the combination of overall and differential attrition rates exceeds WWC standards for this area, and the estimates of effects did not account for the existing differences in pre-intervention characteristics. The high-achieving subgroup did not meet standards because the combination of overall and differential attrition rates exceeds WWC standards for this area, and the subsequent analytic intervention and comparison groups are not equivalent at baseline.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The *Peer-Assisted Learning Strategies* group mean outcome values for Saenz, Fuchs, and Fuchs (2005) were calculated by the WWC using a difference-in-differences approach (see WWC 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.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The mean difference for each of the reading achievement outcomes reported by Saenz, Fuchs, and Fuchs (2005) reflects the mean difference between treatment and control groups calculated by the WWC using the difference-in-differences approach.
5. For an explanation of the effect size calculation, see WWC Procedures and Standards Handbook, Appendix B.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results for the intervention group.
8. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Saenz, Fuchs, and Fuchs (2005), corrections for clustering and multiple comparisons were needed, so the significance levels may differ from those reported in the original study.

Appendix A5 *Peer-Assisted Learning Strategies* rating for the reading achievement domain

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

For the outcome domain of reading achievement, the WWC rated *Peer-Assisted Learning Strategies* for English language learners as having potentially positive effects. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered, as *Peer-Assisted Learning Strategies* 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. *Peer-Assisted Learning Strategies* has one study that shows a substantively important positive effect on reading achievement.

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. *Peer-Assisted Learning Strategies* does not have any studies showing a statistically significant or substantively important negative effect or any studies showing indeterminate effects on reading achievement.

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. *Peer-Assisted Learning Strategies* has no studies showing statistically significant positive effects on reading achievement.

AND

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

  Met. *Peer-Assisted Learning Strategies* has no studies showing statistically significant or substantively important negative effects on reading achievement.

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

Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence[1]
		Schools	Students	
Reading achievement	1	not available	99	Small
English language development	0	na	na	na
Mathematics achievement	0	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.