What Works Clearinghouse

Enhanced Proactive Reading

Program description

Enhanced Proactive Reading, a comprehensive, integrated reading, language arts, and English language development curriculum, is targeted to first-grade English language learners experiencing problems with learning to read through conventional instruction. The curriculum is implemented as small group daily reading instruction, during which English Language Learners instructors provide opportunities for participation from all students and give feedback for student responses.

Research

One randomized controlled trial of *Enhanced Proactive Reading* met the What Works Clearinghouse (WWC) evidence standards and a second randomized controlled trial study met WWC standards with reservations.[1] The two studies included a total of more than 130 students from four schools in Texas. The studies examined results on reading achievement and English language development.[2]

Effectiveness

Enhanced Proactive Reading was found to have potentially positive effects on reading achievement and no discernible effects on English language development.

	Reading achievement	*Mathematics achievement*	*English language development*
Rating of effectiveness	Potentially positive effects	Not reported	No discernible effects
Improvement index[3]	Average: +19 percentile points Range: +2 to +43 percentile points	Not reported	Average: –1 percentile points Range: –7 to +5 percentile points

1. The authors report 10 cases of failed assignment due to scheduling conflicts, the implication of which is the study meets WWC standards with reservations.
2. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.
3. These numbers show the average and range of improvement indices for all findings across the studies.

Additional program information

Developer and contact

Vaughn Gross Center for Reading and Language Arts,[4] College of Education SZB 228, University of Texas at Austin, 1 University Station D4900, Austin, TX 78712-0365. Email: info@texasreading.org. Web: www.texasreading.org/utcrla/. Telephone: (512) 232-2320.

Scope of use

Enhanced Proactive Reading has been implemented with first-grade students. The developers do not indicate if it is applicable to all elementary English language learners. Additional information on the numbers of students, schools, and districts using the intervention is not available.

Teaching

Enhanced Proactive Reading provides 120 daily lessons, each composed of 6–10 short activities, to be delivered throughout the school year. Daily activities typically include playing word games designed to promote phonemic awareness, practicing letter-sound correspondence for letters or letter combinations, practicing writing letters, and learning the sound of a new letter or letter combination. Each lesson is delivered to small groups of students, lasts approximately 50 minutes (40 minutes for lesson delivery and 10 minutes for oracy practice), and focuses on five content strands: phonemic awareness, letter knowledge, word recognition, connected text fluency, and comprehension strategies. Teachers' instructional practices may include visual aids, gestures, and facial expressions to clarify meaning while teaching vocabulary. Student responses are largely choral, with some individual work. Teachers model new content and monitor students' responses to and progress in the fast-paced lessons.

Cost

No information is available on the cost of the program.

Research

Two studies reviewed by the WWC investigated the effects of Enhanced Proactive Reading. One study (Vaughn, Cirino, et al., 2006[5]) was a randomized control trial that met WWC evidence standards. The Vaughn, Cirino, et al. (2006) study included 91 students from four schools. The second study (Vaughn, Mathes, et al., 2006) was a randomized controlled trial that met WWC evidence standards with reservations. The Vaughn, Mathes, et al. study included 41 students from four schools.

In the Vaughn, Cirino, et al. (2006) study 91 Hispanic, first-grade English language learners from Texas received the program as a supplement to their regular reading instruction.[6] All participants were prescreened to confirm scores below the 25th percentile for first grade on the Letter Word Identification subtest of the Woodcock Language Proficiency Battery. A variety of measures of prereading skills and reading levels were administered.

Vaughn, Mathes, et al. (2006) involved 41 first-grade Hispanic English language learners from 14 classrooms in Texas. The

4. The reading and phonological components of the intervention are equivalent to another intervention named Proactive Reading. However, *Enhanced Proactive Reading* includes additional activities geared for English language learners in the areas of language and vocabulary development. Proactive Reading was developed by Mathes and colleagues (2005).
5. Following format conventions (first author and date) would give both of these studies identical names (Vaughn, et al., 2006). This report lists the first two authors' names followed by the date to distinguish between the studies.
6. Although random assignment was done at the student level and the unit of assignment matched the unit of analysis, the authors investigated the magnitude of classroom-level clustering, and clustering that might have occurred because the intervention was delivered in small groups. In both cases, they found clustering had no important impacts on their findings and analyses are presented using student-level findings.

Research *(continued)*

students were randomly assigned to either the treatment or control group.[7] Participants were included in the study based on low English and Spanish reading achievement. The program was used as a supplement to the regular reading program. Individual findings in the reading achievement domain were statistically significant and substantively important. The English language development domain had no statistically significant or substantively important findings.

Effectiveness

Findings

The WWC review of English language learner interventions addresses student outcomes in three domains: reading achievement, mathematics achievement, and English language development.[8]

Reading achievement. Vaughn, Cirino, et al. (2006) found no statistically significant effects for any reading achievement measures. However, five of the seven effect sizes, as well as the average effect size, were large enough to be considered substantively important using WWC criteria. Vaughn, Mathes, et al. (2006) found statistically significant differences favoring the *Enhanced Proactive Reading* students in two of the four reading measures (Word Attack and Passage Comprehension). Further, the average effect size for reading achievement in this study was statistically significant, as calculated by the WWC, and large enough to be considered substantively important using WWC criteria.

English language development. Neither study found a statistically significant or substantively important effect on English language development.

Rating of effectiveness

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

The WWC found *Enhanced Proactive Reading* to have potentially positive effects on reading achievement and no discernible effects on English language development

Improvement index

The WWC computes an improvement index for each individual finding. In addition, within each outcome domain, the WWC computes an average improvement index for each study as well as an average improvement index across studies (see Technical Details of WWC-Conducted Computations). The improvement index represents the difference between the percentile rank of the average student in the intervention condition versus the percentile rank of the average student in the comparison condition. Unlike the rating of effectiveness, the improvement index is entirely based on the size of the effect, regardless of the statistical significance of the effect, the study design, or the analysis. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results. The average improvement index for the reading achievement domain is +19 percentile points, with a range of +2 to +43 percentile

7. The authors report 10 cases of failed assignment due to scheduling conflicts. These students were replaced with alternates prior to onset of the intervention. There was also differential attrition between the two groups, but the authors were able to provide evidence of pretest equivalence for the post-attrition samples.

8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate the statistical significance. In the case of *Enhanced Proactive Reading*, a correction for multiple comparisons was needed.

The WWC found *Enhanced Proactive Reading* to have potentially positive effects on reading achievement and no discernible effects on English language development *(continued)*

points. The average improvement index for the English language development domain is –1 percentile points, with a range of –7 to +5 percentile points.

Summary

The WWC reviewed two studies on *Enhanced Proactive Reading,* one that met WWC evidence standards and one that met WWC evidence standards with reservations. The studies showed potentially positive effects in the reading achievement domain and no discernible effects for the English language development domain. English language learning, an evolving field, is working to establish a research base. The evidence presented in this report is limited and may change as new research emerges.

References

Met WWC evidence standards

Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Cardenas-Hagan, E., Pollard-Durodola, S. D., Fletcher, J. M., & Francis, D. J. (2006). Effectiveness of a Spanish Intervention and an English Intervention for English Language Learners at Risk for Reading Problems. *American Educational Research Journal, 43*(3), 449–487.

Met WWC evidence standards with reservations

Vaughn, S., Mathes, P., Linan-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S., Cardenas-Hagan, E., & Franics, D. (2006). Effectiveness of an English intervention for first-grade English language learners at risk for reading problems. *Elementary School Journal, 107*(2), 153–180.[9]

Additional source:

Mathes, P. G., Denton, C. A., Fletcher, J. M., Anthony, J. L., Francis, D. J., & Schnatschneider, C. (2005). The effects of theoretically different instruction and student characteristics on the skills of struggling readers. *Reading Research Quarterly, 40*, 148–182.

For more information about specific studies and WWC calculations, please see the *WWC Enhanced Proactive Reading Technical Appendices*.

9. This article will be published in the November 2006 issue of *Elementary School Journal.*

Appendix

Appendix A1.1 Study characteristics: Vaughn, Cirino, et al., 2006[1] (randomized controlled trial)

Characteristic	Description
Study citation	Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Cardenas-Hagan, E., Pollard-Durodola, S. D., Fletcher, J. M., & Francis, D. J. (2006). Effectiveness of a Spanish Intervention and an English Intervention for English Language Learners at Risk for Reading Problems. *American Educational Research Journal, 43*(3), 449–487.
Participants	Ninety-one Hispanic, first-grade English language learners (46% female) from 20 classrooms participating in this study were randomly assigned to the intervention or comparison group. All participants were prescreened to assess their English language ability and met two inclusion criteria: scoring below the 25th percentile for first grade on the Letter Word Identification subtest of the Woodcock Language Proficiency Battery; reading between zero and one word from a list of five two- to four-letter words in English and Spanish.
Setting	Four Texas schools, considered effective for bilingual students and where the population was at least 60% Hispanic, were chosen as sites for this study. Schools were located in urban or urban boundary areas. Each school's first-grade classrooms provided core English reading instruction to English language learners using well known reading programs.
Intervention	The intervention group received *Enhanced Proactive Reading* from October to May. The curriculum was implemented as a supplemental reading program for low-performing students. Instruction was carried out over 120 lessons, in 50-minute sessions, to groups of three to five students (with homogeneous reading achievement). Classroom teachers taught both intervention and comparison students. The intervention was delivered by bilingual (Spanish/English) teachers in a pull-out setting. Checklists and observations were conducted and found that the intervention was delivered with acceptable fidelity.
Comparison	The comparison group English language learning students did not participate in the supplemental reading program but received the same core reading instruction as the intervention group. Seventy percent of comparison group students received, on average, 63.6 hours of supplemental reading instruction.
Primary outcomes and measurement[2]	The reading measures included various subtests from the Woodcock-Johnson battery of assessments, the Dynamic Indicators of Basic Early Literacy Skills, and a measure of students' ability to identify letters in the English alphabet and to provide at least one corresponding sound for each letter. The Woodcock Language Proficiency Battery-Revised (composite scores and some related subtests) served as the English language development measure. Although random assignment was done at the student level and the unit of assignment matched the unit of analysis, the study authors investigated classroom-level clustering, and clustering that might have impacted findings because the intervention was delivered in small groups. In both cases, clustering did not have an impact on findings. (See Appendices A2.1 and A2.2 for more detailed descriptions of outcome measures.)
Teacher training	Intervention teachers received 12 hours of professional development prior to implementation and six hours of professional development after the intervention had been implemented for six weeks. Teachers also participated in frequent staff development sessions and on-site coaching. Note that some teachers had previously taught the curriculum as part of the other study reviewed in this report.

1. Following format conventions (first author and date) would give both of the studies reviewed identical names (Vaughn, et al., 2006). This report lists the first two authors' names followed by the date to distinguish between the studies: Vaughn, Cirino et al. & Vaughn, Mathes et al.
2. Pre- and post-intervention assessments included reading assessments in English and Spanish and measures of English language development. Only the English language measures are within the scope of this review and reported here.

Appendix A1.2 Study characteristics: Vaughn, Mathes, et al., 2006 (randomized controlled trial with randomization problems[1])

Characteristic	Description
Study citation	Vaughn, S., Mathes, P., Linan-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S., Cardenas-Hagan, E., & Franics, D. (2006). Effectiveness of an English intervention for first-grade English language learners at risk for reading problems. *Elementary School Journal, 107*(2), 153–180.[2]
Participants	The study involved 41 first-grade Hispanic English language learners (50% female) from 14 classrooms. The students were randomly assigned[2] to either the intervention or comparison group. Participants were included in the study based on low English and Spanish reading achievement. All participants were prescreened to assess their English and Spanish reading and language ability. To be included in the study, students had to meet two inclusion criteria: scoring below the 25th percentile for first grade on the Letter Word Identification subtest of the Woodcock Language Proficiency Battery in both Spanish and English and reading between zero and one word from a list of five two- to four-letter words in English and Spanish.
Setting	Four Texas schools with a large population of English language learners served as sites for the study. On average, 98% of the students were Hispanic and more than 80% (ranging from 85% to 100%) qualified for the free or reduced-price lunch program. Schools were located in large urban areas or on an urban border.
Intervention	The intervention group received *Enhanced Proactive Reading.* The curriculum was implemented as a supplemental reading program for these students. They were taught in small groups of three to five students from October to May, receiving a total of 120, 50-minute lessons. Checklists and observations were conducted and found that the intervention was delivered with acceptable fidelity.
Comparison	The comparison group English language learners did not participate in the supplemental reading program but received the same core reading instruction as the intervention group. The study authors note that 14 of these students received one or more reading interventions, in addition to their core instruction, during the study. More specifically, comparison students received an average of 63.7 hours of supplemental instruction, but there was wide variability, with some receiving relatively few hours and others receiving well over 100 hours (compared with the 80 hours received by students in the intervention group).
Primary outcomes and measurement	Pre- and post-intervention assessments included measures of English language development and reading assessments in both English and Spanish. Only the English language measures are within the scope of this review, and the WWC did not consider Spanish outcomes when assessing the intervention. The reading measures included various subtests from the Woodcock-Johnson battery of assessments, the Dynamic Indicators of Basic Early Literacy Skills, and a measure of students' ability to identify letters in the English alphabet and to provide at least one corresponding sound for each letter. The Woodcock Language Proficiency Battery-Revised served as the English language development measure (composite scores and some related subtests). (See Appendices A2.1 and A2.2 for more detailed descriptions of outcome measures.)
Teacher training	Teachers received 12 hours of professional development training prior to implementation of the intervention and six additional hours six weeks after the start of the intervention. Teachers also participated in frequent staff development sessions and on-site coaching.

1. The authors report 10 cases of failed assignment due to scheduling conflicts, the implication of which is the study meets WWC standards with reservations.
2. This article will be published in the November 2006 issue of *Elementary School Journal.*

Appendix A2.1 Outcome measures[1] in the reading achievement domain

Outcome measure	Description
Letter Sound Identification	Letter Naming and Letter Sound Identification. Students were asked to identify all 26 letters in the English alphabet and to provide at least one sound for each letter. Raw scores were based on the number of correctly named letters and the number of letters where an appropriate sound was provided. The internal consistency of this measure for first-grade bilingual students ranged from 0.94 to 0.97 (as cited in Vaughn, Cirino et al., 2006, and Vaughn, Mathes et al., 2006).
Word Attack	Woodcock Language Proficiency Battery: Word Attack subtest. This is a standardized measure that assesses the student's phonemic awareness skills. Students read a list of nonsense words. Results are available in raw scores, standard scores, Normal Curve Equivalent scores (NCES), age equivalencies, and grade-level equivalencies (as cited in Vaughn, Mathes et al., 2006, and Vaughn, Cirino et al., 2006).
Passage Comprehension	Woodcock Language Proficiency Battery: Passage Comprehension subtest. This is a standardized measure. Students silently read a short passage and then fill in the missing word. Scores are available in raw scores, standard scores, Normal Curve Equivalent scores (NCES), age equivalencies, and grade-level equivalencies (as cited in Vaughn, Cirino et al., 2006, and Vaughn, Mathes et al., 2006)
DIBELS1 & DIBELS2	Dynamic Indicators of Basic Early Literacy Skills: Oral reading fluency & accuracy measure (DIBELS).This measure requires students to read a passage geared at the student's grade level. Children are scored on the number of words read correctly minus the number of words read incorrectly within the time limit. In both studies, two versions of a DIBELS measure were used (passage1 and passage 2). Both passages were designed to be of comparable difficulty (as cited in Vaughn, Cirino et al., 2006, and Vaughn, Mathes et al., 2006[2]).
Word Reading Efficiency	Test of Word Reading Efficiency: Word Reading Efficiency subtest. This test requires students to read isolated words of increasing difficulty in 45 seconds. Students are scored on the number of correct words they read during the allotted time period (as cited in Vaughn, Cirino et al., 2006).
Letter-Word Identification	Woodcock Language Proficiency Battery: Letter-Word Identification subtest.[3] This is a standardized measure that assesses the student's phonemic awareness skills. Students first identify a list of letters and then read a list of words. Scores are available in raw scores, standard scores, Normal Curve Equivalent scores (NCES), age equivalencies, and grade-level equivalencies (as cited in Vaughn, Cirino et al., 2006).

1. Some outcome measures reported by the authors, in both studies, are not reported here because they do not apply to the domain.
2. DIBELS2 outcomes for Vaughn, Mathes et al. (2006) are not reported here. Although posttest scores are available, pretest scores are not. For reasons outlined in Appendix A3.1, the WWC reports gain score effect sizes (that is, pre- to post-change in the treatment group versus the comparison group) for this study.
3. This measure was used for screening purposes (not as an outcome measure) in the Vaughn, Mathes, et al. (2006) study

Appendix A2.2 Outcome measures in the English language development domain

Outcome measure	Description
English Language Composite	Woodcock Language Proficiency Battery—Revised, Oral Language Composite. This is a standardized measure of oral language proficiency composed by a series of subtests in picture vocabulary, listening comprehension, and verbal analogies (as cited in Vaughn, Cirino et al., 2006, and Vaughn, Mathes et al., 2006).
Picture Vocabulary	Woodcock Language Proficiency Battery: Picture Vocabulary subtest. This is a standardized measure. This subtest is a measure of expressive language skills. Students are asked to name familiar and unfamiliar pictured objects. Scores are available in raw scores, standard scores, Normal Curve Equivalent scores (NCES), age equivalencies, and grade-level equivalencies (as cited in Vaughn, Cirino et al., 2006, and Vaughn, Mathes et al., 2006).
Listening Comprehension	Woodcock Language Proficiency Battery: Listening Comprehension subtest. This is a standardized measure. Students are asked to listen to a passage and supply the missing word at the end of a statement using oral cloze procedure. Scores are available in raw scores, standard scores, Normal Curve Equivalent scores (NCES), age equivalencies, and grade-level equivalencies (as cited in Vaughn, Cirino et al., 2006, and Vaughn, Mathes et al., 2006).
Verbal Analogies	Woodcock Language Proficiency Battery: Verbal Analogies subtest. This is a standardized measure. Students answer questions about logical relationships that increase in difficulty. Scores are available in raw scores, standard scores, Normal Curve Equivalent scores (NCES), age equivalencies, and grade-level equivalencies (as cited in Vaughn, Cirino et al., 2006, and Vaughn, Mathes et al., 2006).

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

Outcome measure[3]	Study sample	Sample size (students/ schools)	Author's findings from the study: Mean outcome (standard deviation[2]): *Enhanced Proactive Reading* group	Author's findings from the study: Mean outcome (standard deviation[2]): Comparison group	WWC calculations: Mean difference[4] (*Enhanced Proactive Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at α = 0.05)	WWC calculations: Improvement index[7]
Vaughn, Cirino, et al., 2006 (randomized controlled trial)[8]								
Letter Sound Identification	Grade 1	90	23.02 (4.1)	21.45 (4.5)	1.57	0.26	ns	+10
Word Attack	Grade 1	78	99.92 (14.3)	94.40 (11.8)	5.52	0.42	ns	+16
Letter Word Identification	Grade 1	91	89.88 (19.1)	87.42 (18.5)	2.46	0.13	ns	+5
Passage Comprehension	Grade 1	90	86.93 (13.0)	86.13 (13.3)	0.80	0.06	ns	+2
DIBELS1	Grade 1	88	17.07 (17.9)	12.28 (11.6)	4.79	0.32	ns	+13
DIBELS2	Grade 1	88	14.66 (14.8)	11.26 (10.5)	3.40	0.27	ns	+10
Word Reading Efficiency	Grade 1	89	16.93 (10.3)	12.83 (9.5)	4.10	0.41	ns	+16
Average[9] for reading achievement (Vaughn, Cirino, et al., 2006)						.27	ns	+10
Vaughn, Mathes, et al., 2006 (randomized controlled trial with randomization problems)[8]								
Letter Sound Identification	Grade 1	40	9.53 (2.2)	6.64 (6.9)	2.89	0.57	ns	+22
Word Attack	Grade 1	38	43.71 (12.3)	24.25 (12.4)	19.46	1.53	Statistically significant	+43
Passage Comprehension	Grade 1	39	22.82 (13.8)	4.88 (12.4)	17.94	1.32	Statistically significant	+41
DIBELS1	Grade 1	36	20.49 (14.5)	17.75 (23.8)	2.74	0.14	ns	+6

(continued)

Appendix A3.1 Summary of study findings included in the rating for the reading achievement domain[1] *(continued)*

			Author's findings from the study		WWC calculations			
			Mean outcome (standard deviation[2])					
Outcome measure[3]	Study sample	Sample size (students/ schools)	*Enhanced Proactive Reading* group	Comparison group	Mean difference[4] (*Enhanced Proactive Reading* – comparison)	Effect size[5]	Statistical significance[6] (at α = 0.05)	Improvement index[7]
Average[9] for reading achievement (Vaughn, Mathes, et al., 2006)						0.89	Statistically significant	+28
Domain average[9] for reading achievement across all studies						0.49	na	+19

ns = not statistically significant
na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the improvement index.
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. Some outcome measures reported by the authors, in both studies, are not reported here because they do not apply to the domain.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Note that this appendix shows two different types of mean outcomes. For the Vaughn, Cirino et al. (2006) study the WWC presents posttest differences between the treatment group and comparison group. For the Vaughn, Mathes et al. (2006) study, the WWC presents change score mean differences (that is, the pre- to post-difference in the treatment group minus the pre- to post-difference of the comparison group). Change scores were used in Vaughn, Mathes et al. (2006) because there were 10 cases of failed random assignment and severe differential attrition between the intervention and comparison groups.
5. For an explanation of the effect size calculation, please see the Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition 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.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of the two Vaughn et al. studies, a correction for multiple comparisons was needed for findings in the reading achievement domain, so the significance levels may differ from those reported in the original study.
9. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect size.

Appendix A3.2 Summary of study findings included in the rating for the English language development domain[1]

Outcome measure[3]	Study sample	Sample size (students/ schools)	Author's findings from the study: Mean outcome (standard deviation[2]): *Enhanced Proactive Reading* group	Author's findings from the study: Mean outcome (standard deviation[2]): Comparison group	WWC calculations: Mean difference[4] (*Enhanced Proactive Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at α = 0.05)	WWC calculations: Improvement index[7]
Vaughn, Cirino, et al., 2006 (randomized controlled trial)[8]								
English Language Composite	Grade 1	88	55.66 (16.2)	58.66 (18.8)	–3.00	–0.17	ns	–7
Average[9] for English language development (Vaughn, Cirino, et al., 2006)						–0.17	ns	–7
Vaughn, Mathes, et al., 2006 (randomized controlled trial with randomization problems)								
English Language Composite	Grade 1	39	2.23 (16.9)	–0.19 (18.1)	2.42	0.13	ns	+5
Average[9] for English language development (Vaughn, Mathes, et al., 2006)						0.13	ns	+5
Domain average[9] for English language development across all studies						–0.02	na	–1

ns = not statistically significant
na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the improvement index. Subgroup findings from the same studies are not included in these ratings, but are reported in Appendix A4.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. Some outcome measures reported by the authors, in both studies, are not reported here because they do not apply to the domain.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Note that this appendix shows two different types of mean outcomes. For the Vaughn, Cirino et al. (2006) study the WWC presents posttest differences between the treatment group and comparison group. For the Vaughn, Mathes et al. (2006) study, the WWC presents change score mean differences (that is, the pre- to post-difference in the treatment group minus the pre- to post-difference of the comparison group). Change scores were used in Vaughn, Mathes et al. (2006) because there were 10 cases of failed random assignment and severe differential attrition between the intervention and comparison groups.
5. For an explanation of the effect size calculation, please see the Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition 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.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of both Vaughn et al. studies, no corrections for clustering or multiple comparisons were needed for findings in the English language development domain.
9. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A4 Summary of subscale findings for the English language development domain[1]

Outcome measure[3]	Study sample	Sample size (students/ schools)	Author's findings from the study: Mean outcome (standard deviation[2]): *Enhanced Proactive Reading* group	Author's findings from the study: Mean outcome (standard deviation[2]): Comparison group	WWC calculations: Mean difference[4] (*Enhanced Proactive Reading* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at α = 0.05)	WWC calculations: Improvement index[7]
Vaughn, Cirino, et al., 2006 (randomized controlled trial)[8]								
WJ: Picture Vocabulary	Grade 1	88	56.43 (18.8)	59.68 (19.7)	–3.25	–0.17	ns	+07
WJ: Listening Comprehension	Grade 1	91	51.37 (17.2)	55.40 (18.9)	–4.30	–0.22	ns	+09
WJ: Verbal Analogies	Grade 1	91	77.42 (12.0)	78.79 (13.5)	–1.37	–0.11	ns	+04
Vaughn, Mathes, et al., 2006 (randomized controlled trial)								
WJ: Picture Vocabulary	Grade 1	39	8.3 (20.6)	4.38 (19.4)	3.92	0.19	ns	+08
WJ: Listening Comprehension	Grade 1	39	–7.53 (18.3)	–2.05 (17.1)	–5/48	–0.30	ns	–0.12
WJ: Verbal Analogies	Grade 1	37	4.91 (11.2)	1.73 (8.0)	3.18	0.31	ns	+12

ns = not statistically significant

1. This appendix presents subscale findings for measures that fall in the English language development domain. Total scale scores were used for rating purposes and are presented in Appendix A3.2.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes
3. Some outcome measures reported by the authors, in both studies, are not reported here because they do not apply to the domain.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Again, the WWC reported pre- to post-change scores for the Vaughn, Mathes et al. (2006) study.
5. For an explanation of the effect size calculation, please see the Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition 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.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of both Vaughn et al. studies, no correction for clustering was needed.

Appendix A5.1 *Enhanced Proactive Reading* rating for the reading achievement domain

The WWC rates interventions as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of reading achievement, the WWC rated *Enhanced Proactive Reading* as having potentially positive effects. It did not meet the criteria for positive effects because it had only one study that reported a statistically significant positive effect. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *Enhanced Proactive Reading* was assigned the highest applicable rating.

Rating received

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

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

  Met. One study showed statistically significant positive effects, and two studies showed substantively important positive effects.

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

  Met. Neither study showed a statistically significant or substantively important negative effect. The number of studies showing indeterminate effects (zero) was not greater than the number showing substantively important positive effects (two).

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. *Enhanced Proactive Reading* had only one study that met WWC evidence standards for a strong design and showed statistically significant positive effects.

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

  Met. Neither study showed statistically significant or substantively important negative effects.

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

Appendix A5.2 *Enhanced Proactive Reading* rating for the English language development domain

The WWC rates interventions as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of English language development, the WWC rated *Enhanced Proactive Reading* as having no discernible effects. It did not meet the criteria for positive effects because it had only one study, which had no statistically significant positive effects. Further, it did not meet the criteria for other ratings (potentially positive, mixed, potentially negative, or negative effects), because the study did not show statistically significant or substantively important effects, either positive or negative.

Rating received

No discernible effects: No affirmative evidence of effects.

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

  Met. Neither study reported statistically significant or substantively important effects, either positive or negative.

Other ratings considered

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

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

  Not met. Although one of the two studies met WWC evidence standards for a strong design, no statistically significant positive effects were found.

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

  Met. Neither study reported statistically significant or substantively important negative effects.

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

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

  Not met. Neither study reported a statistically significant or substantively important positive effect.

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

  Not met. Both studies showed indeterminate effects.

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

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect. At least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

  Not met. Neither study reported a statistically significant or substantively important effect in this domain.

OR

(continued)

Appendix A5.2 *Enhanced Proactive Reading* rating for the English language development domain *(continued)*

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

  Not met. Neither study reported a statistically significant or substantively important effect, positive or negative.

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

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

  Not met. The WWC analysis found no statistically significant or substantively important negative effects in this domain.

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

  Met. The WWC analysis found no statistically significant or substantively important positive effects in this domain.

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

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which is based on a strong design.

  Not met. Neither of the studies showed statistically significant or substantively important negative effects.

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

  Met. The WWC analysis found no statistically significant or substantively important positive effects in this domain.

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