

What Works Clearinghouse



Read Naturally¹

Program description *Read Naturally* is designed to improve reading fluency using a combination of books, audio-tapes, and computer software. This program includes three main strategies: repeated reading of English text for oral reading fluency development, teacher modeling of story reading, and systematic monitoring of student progress by teachers. Students work at a reading level appropriate for their achievement level, progress through the program at

their own rate, and work, for the most part, on an independent basis. The *Read Naturally* strategy is designed to increase time spent reading by combining teacher modeling, repeated reading, and progress monitoring. Although the program was not originally developed for English language learners (ELL), materials for these students are now available.

Research One study of a modified version of *Read Naturally* met the What Works Clearinghouse (WWC) evidence standards with reservations. This study included 60 ELL elementary school students from five schools in central Texas and examined effects on students' reading achievement.²

The WWC considers the extent of evidence for *Read Naturally* to be small for reading achievement. No studies that met WWC standards with or without reservations addressed mathematics achievement or English language development.

Effectiveness *Read Naturally* was found to have no discernible effects on elementary school ELL students' reading achievement.

	Reading achievement	Mathematics achievement	English language development
Rating of effectiveness	No discernible effects	na	na
Improvement index ³	Average: 0 percentile points Range: -5 to +6 percentile points	na	na

na = not applicable

1. The study on which this report is based added several components to the *Read Naturally* program. The WWC-ELL Principal Investigator determined that the modified version is close enough to the original that this is a reasonable study of *Read Naturally*.
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 study.

Additional program information

Developer and contact

Read Naturally was developed by Candyce Ihnot. *Read Naturally*, 750 S. Plaza Dr. #100, Saint Paul, MN 55120. Web: www.readnaturally.com. Email: info@readnaturally.com. Telephone: (651) 452-4058 or (800) 788-4085. Fax: (651) 452-9204.

Scope of use

The program was first published in 1991. According to the developer, it has been implemented with special education, Title I, and ELL students throughout the U.S.

Teaching

The *Read Naturally* materials come with a teacher's manual that includes the rationale for the program, descriptions of the materials needed to implement the program, instructions for implementing the program, and sample lesson plans for introducing the program to students. As part of the intervention, students practice reading expository passages until they are able to demonstrate improvement in oral language fluency and appropriate phrasing and expression. As discussed in footnote 1

Research

One study (Denton, Anthony, Parker, & Hasbrouck, 2004) reviewed by the WWC investigated the effects of a modified version of *Read Naturally* on an English language learner sample. Although there was random assignment to treatment groups, three students assigned to the control group were reassigned to the treatment group, and vice versa, one week after the study had begun (as requested by the participating schools). Therefore, this study was determined to be a quasi-experimental design that met the WWC evidence standards with reservations. Data from three students in the comparison group were eliminated from the analysis because of exposure to *Read Naturally* in their classroom, and no data were eliminated from analysis in the treatment group. Although this created differential attrition rates between the study groups (10% attrition in the comparison group and 0% attrition in the treatment group), the authors were

able to demonstrate post-attrition equivalence between groups using the pretest. Denton, Anthony, Parker, & Hasbrouck (2004). The study that examined *Read Naturally* included 60 participants. The *Read Naturally* intervention group received English language pull-out tutoring during the school day in addition to their regular English instruction. The control group received only their regular English language pull-out instruction.

above, the Denton et al. (2004) study made some modifications to the program. The study authors included pre-reading activities and post-reading questioning to facilitate comprehension. Tutors identified two vocabulary words for each *Read Naturally* passage. These words were subsequently introduced to students using sentences from the passages where they appeared; follow-up questions were asked to facilitate discussion about the meaning of a given word. These words were also placed on flashcards and reviewed. Tutors also asked students to identify words they did not know prior to using the repeated reading practice, and they then taught those words. In addition, group size varied from 1–4 students, and completed stories were sent home with students to be read with their parents.

Cost

Individual *Read Naturally* materials range in price from \$5 to \$299. The specific needs of the students served will determine the materials needed and the cost of implementation. There are eight levels of *Read Naturally* materials developed specifically for English language learners. The materials for each level cost \$109.

able to demonstrate post-attrition equivalence between groups using the pretest.

Denton, Anthony, Parker, & Hasbrouck (2004). The study that examined *Read Naturally* included 60 participants. The *Read Naturally* intervention group received English language pull-out tutoring during the school day in addition to their regular English instruction. The control group received only their regular English language pull-out instruction.

Extent of evidence

The WWC categorizes the extent of evidence in each domain as small or moderate to large (see the [What Works Clearinghouse Extent of Evidence Categorization Scheme](#)). The extent of evidence takes into account the number of studies and the

Research (continued)

total sample size across the studies that met WWC evidence standards with or without reservations.⁴

The WWC considers the extent of evidence for *Read Naturally* to be small for reading achievement. No studies that met WWC

standards with or without reservations addressed mathematics achievement or English language development.

Effectiveness Findings⁵

The WWC review of interventions for English language learners addresses student outcomes in three domains: reading achievement, mathematics achievement, and English language development.

Reading achievement. Denton and colleagues (2004) reported, and the WWC confirmed, no statistically significant differences between the intervention and comparison groups on students' reading achievement. In addition, the average effect size was small and deemed not substantively important. Therefore, the one study reviewed showed no discernible effects.

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 *Read Naturally* to have no discernible effects on reading achievement.

Improvement index

The WWC computes an improvement index for each individual finding. In addition, within each outcome domain the WWC computes an average improvement index for each study and an average improvement index across studies (see [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 reading achievement is 0 percentile points, with a range of -5 to +6 percentile points across findings.

Summary

The WWC reviewed one study on *Read Naturally*. This study met WWC standards with reservations. This study found no discernible effects on reading achievement. The evidence presented in this report is limited and may change as new research emerges.

4. The Extent of Evidence categorization was developed to tell readers how much evidence was used to determine the intervention rating, focusing on the number and size of studies. Additional factors associated with a related concept, external validity, such as students' demographics and the types of settings in which studies took place, are not taken into account for the categorization.
5. 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 the [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate the statistical significance. In the case of *Read Naturally*, no corrections for clustering or multiple comparisons were needed.

References **Met WWC evidence standards with reservations**

Denton, C. A., Anthony, J. L., Parker, R., & Hasbrouck, J. E. (2004). Effects of two tutoring programs on the English reading development of Spanish-English bilingual students. *The Elementary School Journal*, 104(4), 289–305.

Additional sources:

Denton, C. A. (2000). The efficacy of two English interventions in a bilingual education program. *Dissertation Abstracts International* 61(11), 4325A. (UMI No. 9994233)
Innot, C. (1992). *Read Naturally*. St. Paul, MN: Read Naturally.

For more information about specific studies and WWC calculations, please see the [WWC Read Naturally Technical Appendices](#).

Appendix

Appendix A1 Study characteristics: Denton, Anthony, Parker, & Hasbrouck, 2004 (quasi-experimental design)

Characteristic	Description
Study citation	Denton, C. A., Anthony, J. A., Parker, R., and Hasbrouck J. E. (2004). Effects of two tutoring programs on the English reading development of Spanish-English bilingual students. <i>The Elementary School Journal</i> , 104(4), 289–305.
Participants	The report by Denton and colleagues covers two studies: one investigates the impacts of <i>Read Naturally</i> and the other investigates the impacts of <i>Read Well</i> . Ninety-three Hispanic ELL students (of which 45 were females) who were identified as having difficulty learning to read English participated in one of the two studies. All students who participated in both studies identified Spanish as their first language and were bilingual. The participants were in second through fifth grade (2nd = 22; 3rd = 37; 4th = 28; 5th = 6) and their ages ranged from 7 years to 12 years (average age = 9 years). The 63 students assigned to the <i>Read Naturally</i> study ¹ were randomly assigned to either the treatment or control group. Three students in the control group were exposed to the treatment. These students were not reassigned to the treatment group. Instead, data from these students were removed from analysis. ² Additionally, as requested by the participating schools, three students assigned to the control group were reassigned to the treatment group, and vice versa, one week after the study had begun. This renders the study a quasi-experimental design. The final sample consisted of 60 students (n = 32 in the treatment group and n = 28 in the control group).
Setting	The students attended one of five schools in a Central Texas school district. The district served 13,664 total students, 32% of whom were Hispanic and 56% of whom were identified as economically disadvantaged.
Intervention	The program occurred during pull-out tutoring sessions during the school day when the participants were not receiving their regular English instruction. Students involved with both programs (<i>Read Naturally</i> and <i>Read Well</i>) received an average of 22 tutoring sessions that were 40 minutes in length. The sessions consisted of repeated oral reading of connected text, vocabulary and comprehension instruction, and systematic monitoring of progress within the program. The standard <i>Read Naturally</i> program was modified for use with English language learners by adding and extending activities related to vocabulary, decoding, and comprehension (such as, oral discussions of vocabulary and comprehension and preteaching important or challenging vocabulary in reading passages).
Comparison	The control group received the same regular English education curriculum as the treatment group but did not receive any additional tutoring.
Primary outcomes and measurement	The study measures in the reading achievement domain included a researcher-developed oral reading assessment ³ and three scales from the Woodcock Reading Mastery Tests-Revised: Word Identification, Word Attack, and Reading Comprehension (see Appendix A2 for more detailed descriptions of outcome measures).
Teacher training	Twenty-three undergraduate students studying special education who were enrolled in a class for teaching students with reading difficulties served as tutors. Tutors received training and were supervised by a graduate student experienced in <i>Read Naturally</i> .

1. Students were assigned to one of the two interventions, *Read Well* or *Read Naturally*, based on their pretest scores on the Word Attack subtest of the Woodcock Reading Master Tests-Revised (WRMT-R). Students who scored below the first-grade equivalency (< 1.0) were assigned to the *Read Well* study, and students whose grade equivalency score was higher than first grade (≥ 1.0) were assigned to the *Read Naturally* study.
2. Because data from three students in the comparison group were eliminated from the analysis and no data were eliminated from analysis in the treatment group, there was differential attrition (10% attrition in the comparison group and 0% attrition in the treatment group). The study did, however, demonstrate post-attrition equivalence.
3. Data from the researcher-developed oral reading assessment were not included in the study. Denton and colleagues (2004) stated that “logistical problems, some instances of potentially unreliable administration, and missing data points resulted in data that is invalid for analyses” (p. 296).

Appendix A2 Outcome measures in the reading achievement domain

Outcome measure	Description
Woodcock Reading Mastery Tests-Revised (WRMT-R): Word Identification subtest	The Word Identification subtest assesses basic reading skills by having participants read words presented in a list (as cited in Denton et al., 2004).
WRMT-R: Word Attack subtest	The Word Attack subtest assesses phonemic decoding by having participants read a list of nonsense words aloud (as cited in Denton et al., 2004).
WRMT-R: Passage Comprehension subtest	The Passage Comprehension subtest uses a cloze format that requires participants to read a passage that has an omitted word. The participants are asked to supply an appropriate word to complete the passage that they are reading (as cited in Denton et al., 2004).

Appendix A3 Summary of study findings included in the rating for the reading achievement domain¹

Outcome measure	Study sample	Sample size (students)	Author's findings from the study					
			Mean outcome (standard deviation ²)		WWC calculations			
			<i>Read Naturally</i> group	Comparison group	Mean difference ³ (<i>Read Naturally</i> – comparison)	Effect size ⁴	Statistical significance ⁵ (at $\alpha = 0.05$)	Improvement index ⁶
Denton et al., 2004 (quasi-experimental design) ⁷								
WRMT-R: Word Identification	Grades 2–5	60	1.12 (11.64)	1.75 (9.65)	-0.63	-0.06	ns	0
WRMT-R: Word Attack	Grades 2–5	60	-0.22 (9.37)	0.97 (8.99)	-1.19	-0.13	ns	-5
WRMT-R: Passage Comprehension	Grades 2–5	60	2.13 (7.90)	0.71 (10.26)	1.42	0.16	ns	+6
Domain average⁸ for reading achievement						-0.01	ns	0

ns = not statistically significant

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. Intervention and control group pre- to posttest change scores were used in the study authors' analyses and in the WWC calculations. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, please see [Technical Details of WWC-Conducted Computations](#).
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between -50 and +50, with positive numbers denoting favorable results.
7. 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 *Read Naturally*, no corrections for clustering or multiple comparisons were needed.
8. This row provides the study average, which in this case is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A4 *Read Naturally* 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 *Read Naturally* as having no discernible effects.

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

Other ratings considered

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

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

Not met. No study that met WWC evidence standards showed statistically significant positive effects. Further, there was only one study, and it did not meet WWC standards for a strong design.

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

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

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

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

Not met. One study showed indeterminate effects and no study showed positive effects.

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

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

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

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

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.

(continued)

Appendix A4 *Read Naturally* rating for the reading achievement domain (continued)

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.

Not met. The WWC analysis found no statistically significant or substantively important negative or 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. No study that met WWC evidence standards showed statistically significant negative effects. Further, there was only one study, and it did not meet WWC standards for a strong design.

- 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.

Appendix A5 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence ¹
		Schools	Students	
English language development	0	na	na	na
Mathematics achievement	0	na	na	na
Reading achievement	1	5	60	Small

na = not applicable/not studied

1. A rating of “moderate 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.”