What Works Clearinghouse

Waterford Early Reading Program™

Program description[1]

Waterford Early Reading Program™ is a software-based curriculum for students in Kindergarten through second grade. The curriculum is designed to promote reading, writing, and typing, incorporating literacy skills such as letter mastery, language stories, spelling, basic writing skills, reading and listening development, and comprehension strategies. It can be used as a supplement to the regular reading curriculum. Program materials include classroom lessons and take-home materials in addition to the Waterford software. *Waterford Early Reading Program*™ offers pretest placement and posttest assessments, in addition to ongoing assessments throughout the program.

Research

One study of *Waterford Early Reading Program*™ met the What Works Clearinghouse (WWC) evidence standards with reservations. The single study included more than 70 Kindergarten students from six schools in Ohio.[2] The WWC considers the extent of evidence for *Waterford Early Reading Program*™ to be small for alphabetics and for comprehension. No studies that met WWC evidence standards with or without reservations addressed fluency or general reading achievement.

Effectiveness

Waterford Early Reading Program™ was found to have potentially positive effects on alphabetics and no discernible effects on comprehension.

	Alphabetics	*Fluency*	*Comprehension*	*General reading achievement*
Rating of effectiveness	Potentially positive effects	na	No discernible effects	na
Improvement index[3]	Average: +19 percentile points Range: –26 to +37 percentile points	na	Average: +4 percentile points	na

na = not applicable

1. The descriptive information for this program was obtained from a publicly available document on the program's web site (http://www.pearsondigital.com/pdfs/werp/researchsummary-werp.pdf, downloaded April 2007). The WWC requests developers to review the program description sections for accuracy from their perspective. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review.
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 student-level improvement indices for all findings across the study.

Additional program information[1]

Developer and contact

Developed by Dustin Heuston, Ph.D., at the Waterford Institute, *The Waterford Early Reading Program*™ is distributed by Pearson Digital Learning. Address: 6710 East Camelback Road, Scottsdale, AZ 85251. Email: pdlinfo@pearson.com. Web: http://www.pearsondigital.com/waterford. Telephone: (888) 977-7900.

Scope of use

The program was initially developed in 1990, and the first level of the *Waterford Early Reading Program*™ was launched in 1995. According to the developer, it is currently used in more than 13,000 sites across the United States and serves 350,000 students.

Teaching

Waterford Early Reading Program™ is divided into three levels, each designed for individualized, year-long instruction. The first level is taught in kindergarten and includes print concepts, phonological awareness, and letter recognition. The second level is taught in first grade and includes letter sounds, word recognition, and beginning reading comprehension. The third level is taught in second grade and builds on levels one and two with an emphasis on content meaning of text and fluency in reading. Each level contains hundreds of songs and game-like activities with color graphics, digitized voices, and animation.

The *Waterford Early Reading Program*™ curriculum includes the Waterford software, assessment materials, classroom lessons, homework materials, and classroom posters, as well as student take-home books, CDs, and handouts. On-site training and online "webinars" are available for initial training in addition to a detailed teacher's guide.

Cost

Program materials for the *Waterford Early Reading Program*™ cost approximately $300 per student. Cost of training is not available.

Research

Thirty-six studies reviewed by the WWC investigated the effects of *Waterford Early Reading Program*™. One study (Hecht & Close, 2002) was a quasi-experimental design that met WWC evidence standards with reservations. The remaining 35 studies did not meet WWC evidence screens.

Hecht & Close (2002) included 42 students in four intervention schools and 34 students in two comparison schools. Students in the intervention schools received *Waterford Early Reading Program*™ in addition to their regular curriculum. Students in the comparison schools received no supplement to their regular curriculum.

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

The WWC considers the extent of evidence for *Waterford Early Reading Program*™ to be small for alphabetics and for comprehension. No studies that met WWC evidence standards with or without reservations addressed fluency or general reading achievement.

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 the students' demographics and the types of settings in which studies took place, are not taken into account for the categorization.

Effectiveness

Findings

The WWC review of interventions for beginning reading addresses student outcomes in four domains: alphabetics, fluency, comprehension, and general reading achievement.[5] The study reported here included outcomes for alphabetics and comprehension. Within alphabetics, the studies reviewed cover four constructs: phonological awareness, letter knowledge, print knowledge, and phonics.

Alphabetics. Hecht & Close (2002) examined nine student outcomes in the alphabetics domain: four phonological awareness outcomes (the Elision, Phonemic Blending, Phonemic Segmenting, and Sound Matching subtests of the Comprehensive Test of Phonological Processing (CTOPP)); one letter identification outcome (a letter name knowledge test); one print awareness outcome (Stones—Concepts About Print test); and three phonics outcomes (the Letter Sound Knowledge and Letter Word Identification subtests of the Woodcock-Johnson Tests of Achievement–Revised and the Spelling subtest (with phonemic representation scoring)[6] of the Wide Range Achievement Test (WRAT)). Hecht & Close (2002) reported statistically significant positive effects of the *Waterford Early Reading Program*™ for all nine outcomes. However, the WWC analysis found that none of these effects were statistically significant. The average effect size across all nine outcomes was large enough to be considered substantively important according to the WWC criteria (that is, an effect size of at least 0.25).

Comprehension. The study authors examined one vocabulary development outcome (the Vocabulary subtest of Stanford-Binet, Fourth Edition) and reported no statistically significant effect. The effect size was neither statistically significant nor substantively important.

Rating of effectiveness

The WWC rates the effects of an intervention in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative. The rating of effectiveness takes into account four factors: the quality of the research design, the statistical significance of the findings,[7] the size of the difference between participants in the intervention and the comparison conditions, and the consistency in findings across studies (see the WWC Intervention Rating Scheme).

The WWC found *Waterford Early Reading Program*™ to have potentially positive effects on alphabetics and no discernible effects on comprehension

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 based entirely on the size of the effect, regardless of the statistical significance of the effect, the study design, or the analyses. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.

The average improvement index for alphabetics is +19 percentile points for the single study, with a range of –26 to +37 percentile points across findings. The improvement index for the single outcome in the comprehension domain is +4 percentile points.

5. For definitions of the domains, see the Beginning Reading Protocol.
6. Spelling is not typically counted as a WWC Beginning Reading outcome, but this subtest used a qualitative scoring method. See Appendix A2.1 for details.
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, 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 *Waterford Early Reading Program*™, corrections for clustering and multiple comparisons were needed.

The WWC found *Waterford Early Reading Program*™ to have potentially positive effects on alphabetics and no discernible effects on comprehension *(continued)*

Summary

The WWC reviewed 36 studies on *Waterford Early Reading Program*™. One study met WWC evidence standards with reservations; the remaining studies did not meet WWC evidence screens. Based on the one study, the WWC found potentially positive effects on alphabetics and no discernible effects on comprehension. The evidence presented in this report may change as new research emerges.

References

Met WWC evidence standards with reservations

Hecht, S. A., & Close, L. (2002). Emergent literary skills and training time uniquely predict variability in responses to phonemic awareness training in disadvantaged kindergartners. *Journal of Experimental Child Psychology, 82*(2), 93–115.

Additional citation for this study:

Hecht, S. A. (2000). *Waterford Early Reading program in Ohio: An evaluation.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)

Did not meet WWC evidence screens

Alfaro, R. (1999). The technology-reading connection. *Educational Workshop, 56*(6), 48–51.[8]

Canedo, M., Smolen, L., & Pollard, J. (2000). *A study of the effectiveness of the Waterford Early Reading Program: Final evaluation results 1997–98. Buffalo, NY: Buffalo Public Schools.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[9]

Cassady, J. C., & Smith, L. L. (2003). The impact of a reading-focused integrated learning system on phonological awareness in kindergarten. *Journal of Literacy Research, 35*(4), 947–964.[10]

Cassady, J. C., & Smith, L. L. (2005). The impact of a structured integrated learning system on first grade students' reading gains. *Reading and Writing Quarterly, 21*(4), 361–376.[11]

Cope, R., & Cummings, J. (2001). *Evaluation of the Waterford Early Reading Program in Madisonville Consolidated Independent School District.* Huntsville, TX: Sam Houston State University. (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[12]

Corbett, R. (2000). *An evaluation of the Waterford Early Reading Program in the Hillcrest Title I school in the Alpine School District.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Dunn, A. (1999). *Evaluation of the Waterford Early Reading Program in selected Utah public school for the 1997–98 school year. Salt Lake City, UT: Waterford Institute, Inc.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[11]

Electronic Education. (2002). *Correlation. Put reading first: The research building blocks for teaching children to read.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Heuston, D. (1996). Power tools. *Phi Delta Kappan, 77*(10), 706.[8]

8. Does not use a strong causal design: this study did not use a comparison group.
9. Complete data are not reported: the WWC could not compute effect sizes because complete study details were not reported.
10. Does not use a strong causal design: there was only one intervention and one comparison unit, so the analysis could not separate the effects of the intervention from other factors.
11. Does not use a strong causal design: a historical cohort was used as the comparison group. WWC conventions allow for historical control cohorts in studies that have a broad unit of analysis (school or higher). This study analyzed at the student level and therefore does not fulfill the WWC requirement.
12. Incomparable groups: this study was a quasi-experimental design that used achievement pretests but it did not establish that the comparison group was comparable to the treatment group prior to the start of the intervention.

References *(continued)*

Hillsborough County Public Schools. (n.d.). *An evaluation of the Waterford Early Reading Program Bryan Elementary Hillsborough County, Florida 1997–98 school year.* Plant City, FL: Author. (Available from Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Jadali, F., & Wright, B. (n.d.). *A study of the effectiveness of the Waterford Early Reading Program at Commons Lane Elementary School.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Kortz, W. J. (n.d.). *An evaluation of Waterford Early Reading Program.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101) (Study: Hempstead Independent School District, Hempstead, TX)[8]

Murray-Ward, M. (2000). *El Centrito interim grant report for the period of July 1, 1999 to December 31, 1999 (Report No. 109).* Thousand Oaks: California Lutheran University, Educational Research and Leadership Institute. (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[13]

Nagel, R. J. (1999). *An evaluation of the Waterford Early Reading Program in Pittsburgh, PA Public School District.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[14]

Obeso-Bradley, C., & Miller, B. (1999, December). *Early literacy and technology: The Waterford Early Reading Program (WERP) Level 2, Southside School District, Hollister, California.* Paper presented at the annual education conference of the California School Boards Association, San Francisco, CA. (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Paterson, W. A., Henry, J. J., O'Quin, K., Ceprano, M. A., & Blue, E. V. (2003). Investigating the effectiveness of an integrated learning system on early emergent readers. *Reading Research Quarterly, 38*(2), 172–206.[15]

Pukas, B. (1998, May). *A study of the effectiveness of the Waterford Early Reading Program: First-year evaluation reviews.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101) **(Study: New London Public Schools)**[10]

Research, Assessment & Measurement, Inc. (1997). *A study of the effectiveness of the Waterford Program at Glenridge Elementary School. Baltimore, MD: Author.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[10]

Research, Assessment & Measurement, Inc. (1999). *Evaluation of Waterford Early Reading Program Hacienda la Puente Unified School District Los Angeles County, CA Program Year 1997–98.* Baltimore, MD: Author. (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[12]

Research, Assessment & Measurement, Inc. (2004). *Evaluation of Waterford Early Reading Program: Collins Garden and Nelson Elementary Schools San Antonio, Texas.* Baltimore, MD: Author. (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Reynolds, C. (2001). *An evaluation of the Waterford Early Reading Program.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101) **(Study: Decatur School District 61)**[14]

Roe, E. (2000). *An evaluation of the Waterford Early Reading Program in Scott Lane Elementary School Santa Clara Unified School District, Santa Clara, Calif., 1998–99 school year.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

13. The sample is not appropriate to this review: the parameters for this WWC review specified that students should be in grades kindergarten through third grade during the time of the intervention; this study did not focus on the targeted grades.
14. Does not use a strong causal design: this study was a quasi-experimental design but did not use achievement pretests to establish that the comparison group was equivalent to the intervention group at baseline.
15. Does not use a strong causal design: this study was a quasi-experimental design but did not provide enough information to establish that the comparison group and the intervention group were composed of comparable students.

References *(continued)*

Shapley, K. S. (1997). *Special report of the 1996–1997 Waterford Early Reading Program.* Dallas, TX: Dallas Public Schools.[14]

Additional citation for this study:

Waterford Institute, Inc. (1998). *A preliminary report of the 1996–97 test results from elementary schools in the Dallas ISD on the effectiveness of the Waterford Early Reading Program.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)

Tracey, D. H. (n.d.). *The Waterford Early Reading Program: Research orientation, studies, and findings: Executive summary.* Retrieved from Pittsburgh Public Schools, Office of the Deputy Superintendent of Instruction, Assessment, and Accountability Web site: http://www.pps.k12.pa.us/academicoffice/literacyplus/waterford/stuff/executive%20summary%20-%20diane%20traecy.doc[8]

Walberg, H. J. (2001). *Final evaluation of the reading initiative* (Report to the J. A. & Kathryn Albertson Foundation Board of Directors). Available from the Waterford Institute Web site: http://www.waterford.org/corporate_pages/IdahoStudy.pdf[8]

Washington, S. T. (2003). Teachers' perceptions of the implementation of the Waterford Early Reading Program, a computer-based instruction program: A case study of the evidence from teachers' interviews and students' achievement data in selective Pennsylvania urban elementary schools. *Dissertation Abstracts International, 64*(07), 2341A. (UMI No. 3099994)[9]

Waterford Institute, Inc. (1996). *Preliminary research: Waterford Institute's Early Reading Program—Utah and New York schools.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Waterford Institute, Inc. (1998a). *A study of the correlation between test gains and time spend using the Waterford Early Reading Program.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Waterford Institute, Inc. (1998b). *Preliminary research data on the effect of the Waterford early reading program based on daily use of computer materials for 15 minutes.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101) **(Study: Provo, Utah and New York City)**[8]

Waterford Institute, Inc. (1998c). *Waterford Early Reading Program (Level 1)—Hillcrest Elementary School preliminary study.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Waterford Institute, Inc. (2000). *Evaluation of the Waterford Early Reading Program Level 1 Norwalk Public Schools: Norwalk, CT 1998–99 school year.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[10]

Waterford Institute, Inc. (2002). *Los Angeles Unified School District Waterford Early Reading Program initial implementation findings.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Waterford Institute, Inc. (2004). *A study of the effectiveness of the Waterford Early Reading Program in academic alliance and intensive academic support classrooms in the Los Angeles Unified School District.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[16]

Waterford Institute, Inc. (n.d.). *Waterford Early Reading Program implementation results 1996–97: Duncanville I.S.D.* (Available from the Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101)[8]

Young, J. W., & Tracey, D. H. (2004). *An evaluation of the Waterford Early Reading Program Newark, New Jersey 1997–98 school year.* Sandy, UT: Waterford Institute, Inc.[17]

For more information about specific studies and WWC calculations, please see the WWC *Waterford Early Reading Program*™ Technical Appendices.

16. Confound: this study included *Waterford Early Reading Program*™ but combined it with another intervention so the analysis could not separate the effects of the intervention from other factors.
17. Complete data are not reported: the WWC could not evaluate the design because complete data were not reported. Attempts to contact the authors for more information were unsuccessful.

Appendix

Appendix A1 Study characteristics: Hecht & Close, 2002 (quasi-experimental design)

Characteristic	Description
Study citation	Hecht, S. A., & Close, L. (2002). Emergent Literacy Skills and Training Time Uniquely Predict Variability in Responses to Phonemic Awareness Training in Disadvantaged Kindergartners. *Journal of Experimental Child Psychology*, *82*(2), 93–115. *Additional source:* Hecht, S. A. (2000). *Research Compendium: The Waterford Early Reading Program.* (Available from Waterford Institute, Inc., 55 West 900 South, Salt Lake City, UT 84101). (Study: Waterford Early Reading program in Ohio.)
Participants	The study began with 140 full-day, at-risk Kindergarten students who were randomly selected from six schools. Students from four schools who received the *Waterford Early Reading Program*™ were matched to students in two schools who did not receive the program. Students were pretested in the fall and posttested in the spring of the same school year. Because of mobility and absences, 64 students attrited from the study. The final analysis sample included 76 students. The mean age of students was five years and seven months. The majority of students were eligible to receive free/reduced lunch.[1] The majority of students in the schools came from low socio-economic status and African-American families.
Setting	The study took place in six inner city or rural public schools in Ohio.
Intervention	Students received the computer-assisted instruction of *Waterford Early Reading Program*™–Level One (WERP–1) during their normal classroom lessons for six months. The program focused on phonological awareness skills, letter knowledge, print concepts, and oral language skills. Students worked on the Waterford multimedia computer on their own for 15 minutes each session. A teacher management system was used to track daily time use.
Comparison	Students in the comparison group received their regular reading curriculum and were not exposed to the *Waterford Early Reading Program*™.
Primary outcomes and measurement	Nine outcomes were assessed in the alphabetics domain including the Comprehensive Test of Phonological Process (Phonemic Segmenting, Phonemic Blending, Elision, and Sound Matching subtests), the Woodcock-Johnson Tests of Achievement (the Letter Word Identification subtest), the Wide Range Achievement Test (the Spelling subtest with Phonemic Representation scoring), the Concepts About Print Test, the Letter Name Knowledge and Letter Sound Knowledge measures, and the Stanford-Binet: Fourth Edition Vocabulary subtest. The study also used a letter writing task from the Spelling subtest of the Wide Range Achievement Test, but this test was outside the domains specified by the Beginning Reading protocol (see Appendices A2.1–2.2 for more detailed descriptions of outcome measures).
Teacher training	Information about teacher training was not provided in the study.

1. The WWC received additional information on the analytic sample from the study authors. Baseline equivalence of the intervention and comparison group students remaining in the study was demonstrated by the authors.

Appendix A2.1 Outcome measures in the alphabetics domain

Outcome measure	Description
Phonological awareness	
Comprehensive Test of Phonological Processing (CTOPP): Elision subtest	A standardized measure of children's phonological awareness skills. Children were asked to say a word. Then children were asked what the word would be if a specific phoneme in the word were deleted. The remaining phonemes were used to form a word (as cited in Hecht, 2000).
CTOPP: Phonemic Blending subtest	A standardized measure of children's phonemic synthesis skills. It includes four practice items and 15 test items consisting of two- to four-phoneme, one- and two-syllable words. This test measures the total number of words correctly spoken (as cited in Hecht & Close, 2002).
CTOPP: Phonemic Segmenting subtest	A standardized measure of children's phonemic analysis skills. It includes three practice items and 15 test items consisting of two- to five-phoneme single-syllable words. This test measures the total number of words correctly pronounced one phoneme at a time (as cited in Hecht & Close, 2002).
CTOPP: Sound Matching subtest	A standardized measure of children's sound matching skills. Children were asked to pick which of three pictured words began with the same first sound as a target word (as cited in Hecht, 2000).
Letter identification	
Letter Name Knowledge	A researcher-developed measure designed to measure the total number of letter names correctly pronounced (as cited in Hecht & Close, 2002).
Print awareness	
Concepts About Print Test	This 18-question test (Stones version) yielded one score reflecting students' knowledge about print. The score is measured by the total number of correct items (as cited in Hecht & Close, 2002).
Phonics	
Letter Sound Knowledge	A researcher-developed measure designed to measure the total number of letter sounds correctly pronounced (as cited in Hecht & Close, 2002).
Wide Range Achievement Test: Spelling subtest with phonemic representation scoring	Students wrote 15 words as dictated by the test administrator. Scoring was based on Wilkinson's method of giving partial credit for accuracy of phonemic representation (as cited in Hecht & Close, 2002). Students received between 0 and 6 points depending on how many and the placement of phonemes were represented by phonemically related or conventional letters in each written word.
Woodcock-Johnson Tests of Achievement-Revised: Letter Word Identification subtest	A standardized measure of children's word reading. Children identified various letters of the alphabet as well as words, ranging from commonly used words to less familiar words of the English language (as cited in Hecht, 2000).

Appendix A2.2 Outcome measure in the comprehension domain

Outcome measure	Description
Vocabulary	
Stanford-Binet (4th ed.): Vocabulary subtest	A standardized measure to assess general cognitive ability and estimate general verbal IQ. The score is measured by the total number of correctly defined words (as cited in Hecht & Close, 2002).

Appendix A3.1 Summary of study findings included in the rating for the alphabetics domain by construct[1]

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation[2]) — *Waterford* group	Authors' findings from the study: Mean outcome (standard deviation[2]) — Comparison group	WWC calculations: Mean difference[3] (*Waterford* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Hecht & Close, 2002 (quasi-experimental design)[7]								
Phonological awareness								
CTOPP: Elision subtest	Kindergarten	6/76	4.71 (3.45)	2.82 (2.39)	1.89	0.62	ns	+23
CTOPP: Phonemic Blending subtest	Kindergarten	6/76	9.53 (5.55)	4.24 (5.08)	5.29	0.98	ns	+34
CTOPP: Phonemic Segmenting subtest	Kindergarten	6/76	7.58 (7.05)	1.53 (2.84)	6.05	1.07	ns	+36
CTOPP: Sound Matching subtest	Kindergarten	6/76	10.91 (4.71)	6.27 (4.89)	4.64	0.96	ns	+33
Letter identification								
Letter Name Knowledge	Kindergarten	6/76	21.58 (4.43)	24.65 (4.14)	–3.07	–0.71	ns	–26
Print awareness								
Concepts About Print Test	Kindergarten	6/76	8.58 (3.05)	9.01 (4.57)	–0.43	–0.11	ns	–4
Phonics								
Letter Sound Knowledge	Kindergarten	6/76	19.09 (8.87)	22.55 (9.33)	–3.46	–0.38	ns	–15
WRAT: Spelling subtest with phonemic representation scoring	Kindergarten	6/76	25.57 (19.67)	8.09 (7.79)	17.48	1.11	ns	+37
Woodcock-Johnson Tests of Achievement-Revised: Letter Word Identification subtest	Kindergarten	6/76	3.54 (3.43)	0.77 (1.16)	2.67	0.99	ns	+34
Domain average[8] for alphabetics						0.50	ns	+19

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for alphabetics.

(continued)

Appendix A3.1 Summary of study findings included in the rating for the alphabetics domain by construct *(continued)*

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. The standard deviations for CTOPP Elision and Sound Matching subtests were received from the first author.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The intervention group mean in this table equals the comparison group mean plus the mean difference. The mean difference is calculated as the difference between gain scores and takes into account the pretest difference between the study groups.
4. For an explanation of the effect size calculation, 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 versus the percentile rank of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.
7. The 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 Hecht & Close (2002), corrections for clustering and multiple comparisons were needed, so the significance levels differ from those reported in the original study.
8. This row provides the study average, which in this instance, is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

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

Outcome measure	Study sample	Sample size (schools/ students)	Authors' findings from the study: Mean outcome (standard deviation[2]) — *Waterford* group	Authors' findings from the study: Mean outcome (standard deviation[2]) — Comparison group	WWC calculations: Mean difference[3] (*Waterford* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Hecht & Close, 2002 (quasi-experimental design)								
Stanford-Binet (4th ed.): Vocabulary subtest	Kindergarten	6/76	16.91 (3.66)	16.58 (3.35)	0.33	0.09	ns	+4

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for comprehension.
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. The intervention group mean in this table equals the comparison group mean plus the mean difference. The mean difference is calculated as difference between gain scores and takes into account the pretest difference between the study groups.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see 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 versus the percentile rank of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.

Appendix A4.1 *Waterford Early Reading Program*™ rating for the alphabetics domain

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

For the outcome domain of alphabetics, the WWC rated *Waterford Early Reading Program*™ as having potentially positive effects. It did not meet the criteria for positive effects because no studies showed statistically significant positive effects. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, negative effects) were not considered because the intervention 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. The single study of *Waterford Early Reading Program*™ showed substantively important positive effects.

AND

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

  Met. No studies of *Waterford Early Reading Program*™ showed statistically significant or substantively important negative effects, and no studies showed indeterminate effects.

Other ratings considered

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

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

  Not met. No studies of *Waterford Early Reading Program*™ showed statistically significant positive effects.

AND

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

  Met. The single study of *Waterford Early Reading Program*™ did not show statistically significant or substantively important negative effects.

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

Appendix A4.2 *Waterford Early Reading Program*™ rating for the comprehension domain

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

For the outcome domain of comprehension, the WWC rated *Waterford Early Reading Program*™ as having no discernible effects. It did not meet the criteria for other ratings (positive effects, potentially positive effects, mixed effects, potentially negative effects, and negative effects) because the one study that met WWC standards with reservations did not show statistically significant or substantively important 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 one study of *Waterford Early Reading Program*™ did not show a statistically significant or substantively important effect, either positive or negative.

Other ratings considered

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

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

  Not met. The one study of *Waterford Early Reading Program*™ did not show a statistically significant positive effect.

AND

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

  Met. The one study of *Waterford Early Reading Program*™ did not show a statistically significant or substantively important negative effect.

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

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

  Not met. The one study of *Waterford Early Reading Program*™ did not show a statistically significant or substantively important positive effect.

AND

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

  Met. The one study of *Waterford Early Reading Program*™ did not show a statistically significant or substantively important negative effect. This study showed an indeterminate effect.

(continued)

Appendix A4.2 *Waterford Early Reading Program*™ rating for the comprehension domain *(continued)*

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 one study of *Waterford Early Reading Program*™ did not show statistically significant or substantively important effects.

OR

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

  Not met. No studies showed statistically significant or substantively important effects.

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.

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

AND

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

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

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

  Not met. No studies showed statistically significant negative effects.

AND

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

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

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

Appendix A5 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence[1]
		Schools	Students	
Alphabetics	1	6	76	Small
Fluency	0	0	0	na
Comprehension	1	6	76	Small
General reading achievement	0	0	0	na

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."