Appendix

Appendix A1.1 Study characteristics: Borman & Benson, 2006 (randomized controlled trial)

Characteristic	Description
Study citation	Borman, G. D., & Benson, J. (2006). *Can brain research and computers improve literacy? A randomized field trial of the Fast ForWord® Language computer-based training program* (WCER Working Paper No. 2006-5). Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research.
Participants	Students were eligible for the study if they scored below national norms on the total reading outcome for the district-administered Comprehensive Test of Basic Skills, Fifth Edition (CTBS/5) during the spring of 2000. One hundred and forty-one academically at-risk second-grade students took pretests (CTBS/5 test) in the spring of 2001.[1] Random assignment was conducted separately within each school. Of the initial intervention (71) and comparison (70) students, listwise deletion of students with missing pretest or posttest data was conducted. This resulted in an analysis sample of 60 intervention and 52 comparison students. Additionally, six students (two from the intervention group and four from the comparison group) were dropped from the sample because they were determined to be outliers based on a substantial drop from pre- to posttest. The groups primarily consisted of African-American (92% of the intervention and 94% of the comparison) and economically disadvantaged students (75% of both groups received free lunch). There were slightly more male participants (52% of the intervention and 56% of the comparison).
Setting	The study took place in four urban schools in the Baltimore City Public School System.
Intervention	In addition to their regular reading instruction, students randomly assigned to the intervention condition used the *Fast ForWord® Language* software program in school resource rooms. The resource rooms served as a targeted pullout program offered during the regular school day supplementing the regular classroom literacy instruction. Students received the program 100 minutes a day, five days a week, for at least 20 days between April and June 2001, under the supervision of a *Fast ForWord®*–trained teacher.
Comparison	In addition to their regular reading instruction, comparison group students received non-literacy instruction or participated in special activities and classes, such as art and gym.
Primary outcomes and measurement	The total reading portion of the CTBS/5 was used as an outcome measure. The Terra Nova was used as both the pretest (Form B in April 2001) and the posttest (Form A in June 2001).[2]
Teacher training	Before the start of the program, Scientific Learning provided training sessions for teachers operating the *Fast ForWord®* programs at the schools.

1. The study also included 274 seventh-grade students in elementary/middle and middle schools, but these students do not fall within the scope of the WWC's Beginning Reading review.
2. The study also included CTBS Language scores, but this measure does not fall within the scope of the WWC's Beginning Reading review.

Appendix A1.2 Study characteristics: Scientific Learning Corporation, 2005a (randomized controlled trial)

Characteristic	Description
Study citation	Scientific Learning Corporation. (2005a). Improved early reading skills by students in three districts who used Fast ForWord® to Reading 1. *MAPS for Learning: Product Reports, 9*(1), 1–5.
Participants	During the spring of the 2004/05 school year, 158 first- and 50 second-grade students from three different schools participated in the study. At one school, students from both grades participated, while only first-grade students participated at the other two schools. Using random assignment within school and grade, 103 low-achieving students were assigned to the *Fast ForWord®* group (78 first- and 25 second-grade students) and 105 students served as a comparison group (80 first- and 25 second-grade students). Four students (two intervention and two comparison) were older than nine years at one or both testing times, which is too old for the norms of the Test of Phonological Awareness (TOPA), so they were removed from the analysis sample. Additionally, three intervention and four comparison students moved during the study. Therefore, the analysis sample included 197 students: 75 first- and 23 second-grade students in the intervention group and 78 first- and 21 second-grade students in the comparison group. Seven study participants (one intervention and six comparison students) had used the *Fast ForWord® Basics* product before participating in the study.
Setting	The three schools were located in different districts in different states, including one rural and one urban district. Results for one of the districts, Springfield City School District in Springfield, Ohio, was presented in a separate manuscript and can be viewed in Appendix A4.
Intervention	All students in the *Fast ForWord®* group used the *Fast ForWord® to Reading 1* product, a computer-based product designed using first-grade curriculum standards. The *Fast ForWord® to Reading 1* protocol called for students to use the product for 48 minutes a day, five days a week, for eight to 12 weeks. Students were pulled out of class to use the program in a computer lab, where two paraprofessionals monitored the students, but did not assist with the content except to give instructions.
Comparison	Students in the comparison group took part in the regular school curriculum.
Primary outcomes and measurement	The phonological awareness and letter-sounds subtests of the early elementary version of the TOPA were used for both the pre- and posttest.
Teacher training	The intervention teachers were given background information on how phonemic awareness and the acoustic properties of speech can impact development of language and reading skills. They were then trained to implement the program, including approaches for using Progress Tracker, the program's reporting system, to monitor student performance. Teachers were also trained to assess potential participants for the study and to deliver evaluation outcomes.

Appendix A1.3 Study characteristics: Scientific Learning Corporation, 2005b (randomized controlled trial)

Characteristic	Description
Study citation	Scientific Learning Corporation. (2005b). Improved reading skills by students in the Lancaster County School District who used Fast ForWord® to Reading 2. *MAPS for Learning: Educator Reports, 9*(8), 1–4.
Participants	During the spring of the 2004/05 school year, 50 third-grade students participated in the study, including one entire classroom of students along with randomly selected students from other third-grade classrooms. Using random assignment, 25 students were assigned to the *Fast ForWord®* group and 25 students to a comparison group. All study participants had used one or more of the *Fast ForWord®* products before participating in the study; however, none had previously used *Fast ForWord® to Reading 2*, the focus of this study. Approximately 63% of the students in the study school were Caucasian and 35% were African-American. Nearly 36% of students received free or reduced-price lunch.
Setting	The students attended a K–5 elementary school in Lancaster, South Carolina.
Intervention	All students in the *Fast ForWord®* group used the computer-based *Fast ForWord® to Reading 2* product. The *Fast ForWord® to Reading 2* protocol called for students to use the product for 48 to 90 minutes a day, five days a week, for four to 12 weeks. The entire class of students received the intervention, with a lab manager and a certified teacher talking to each student and discussing what areas needed improvement. Students missed the social studies and science portions of the school curriculum during participation in the intervention.
Comparison	All students were using SRA/McGraw-Hill's Open Court Reading for their whole group reading instruction as part of their regular school curriculum.
Primary outcomes and measurement	The sight word efficiency and phonemic decoding efficiency subtests of the Test of Word Reading Efficiency (TOWRE) were used as both the pretest and the posttest.
Teacher training	The intervention teachers were given background information on how phonemic awareness and the acoustic properties of speech can impact development of language and reading skills. They were then trained to implement the program, including approaches for using Progress Tracker, the program's reporting system, to monitor student performance. Teachers were also trained to assess potential participants for the study and to deliver evaluation outcomes.

Appendix A1.4 Study characteristics: Scientific Learning Corporation, 2005c (randomized controlled trial)

Characteristic	Description
Study citation	Scientific Learning Corporation. (2005c). Improved reading skills by students in Seminole County School District who used Fast ForWord® to Reading 1 and 2. *MAPS for Learning: Educator Reports, 9*(17), 1–6.
Participants	During the spring of the 2004/05 school year, 15 second- and 23 third-grade students—all low-achieving—participated in the study. Using random assignment stratified by grade, academic skill level, and previous *Fast ForWord®* use, 20 students were assigned to the *Fast ForWord®* group and 18 students to a comparison group. Sixteen study participants had used one or more of the *Fast ForWord®* products before participating in the study; however, none had previously used *Fast ForWord® to Reading 1* or *2*, the focus of this study. At the study school, approximately 56% of the students were Caucasian, 22% were of Hispanic origin, and 21% were African-American. Nearly two-thirds of students in the study (compared with 57% of students at the school) were receiving free or reduced-price lunch.
Setting	The students attended an urban pre-K to grade 5 elementary school in Fern Park, Florida.
Intervention	All students in the *Fast ForWord®* group used the *Fast ForWord® to Reading 1* or *2* products. The *Fast ForWord® to Reading 1* and *2* protocols called for students to use the product for 48 minutes a day, five days a week, for four to eight weeks. Students were pulled out of class to use the program in a computer lab, where a certified teacher and paraprofessional oriented the students on the product and made sure they understood the tasks. Once the students started the product, no assistance was given.
Comparison	Students in the comparison group took part in the regular school curriculum.
Primary outcomes and measurement	The Degrees of Reading Power (DRP) test was used as both the pretest and the posttest.
Teacher training	The intervention teachers were given background information on how phonemic awareness and the acoustic properties of speech can impact development of language and reading skills. They were then trained to implement the program, including approaches for using Progress Tracker, the program's reporting system, to monitor student performance. Teachers were also trained to assess potential participants for the study and deliver evaluation outcomes.

Appendix A1.5 Study characteristics: Scientific Learning Corporation, 2006 (randomized controlled trial)

Characteristic	Description
Study citation	Scientific Learning Corporation. (2006). Improved reading skills by students who used Fast ForWord® to Reading Prep. *MAPS for Learning: Product Reports, 10*(1), 1–6.
Participants	During the fall of the 2005/06 school year, 48 low-performing Kindergarten students participated in the study. Using random assignment, 25 students were assigned to the *Fast ForWord®* group and 23 students to a comparison group. A total of seven students in the study were receiving other services: four in the intervention group (one for speech, two for special education, and one was an English language learner) and three in the comparison group (two for speech and one for special education).
Setting	The students attended one suburban elementary school.
Intervention	All students in the *Fast ForWord®* group used the *Fast ForWord® to Reading Prep* product. The *Fast ForWord® to Reading Prep* protocol called for students to use the product for 30 minutes a day, five days a week, for 12 to 16 weeks. Students were pulled out of their classroom at the beginning of the day.
Comparison	Students in the comparison group took part in the regular school curriculum, which included oral language and group activities.
Primary outcomes and measurement	The Dynamic Indicators of Basic Early Literacy Skills (DIBELS) initial sound fluency and letter naming fluency subtests and Woodcock Johnson (WJ) letter word identification subtest were administered as pretests in mid-September and as posttests in mid-December. The Test of Phonological Awareness (TOPA) and Reading Edge (initial sound discrimination, initial sound knowledge, and non-word recognition subtests) were also administered as posttests in mid-December. Findings on the TOPA and Reading Edge tests were not included in the original study, but were provided directly to the WWC by the study authors.
Teacher training	The intervention teachers were given background information on how phonemic awareness and the acoustic properties of speech can impact development of language and reading skills. They were then trained to implement the program, including approaches for using Progress Tracker, the program's reporting system, to monitor student performance. Teachers were also trained to assess potential participants for the study and to deliver evaluation outcomes.

Appendix A1.6 Study characteristics: Overbay & Baenen, 2003 (quasi-experimental design)

Characteristic	Description
Study citation	Overbay, A., & Baenen, N. (2003). *Fast ForWord® Evaluation, 2002–03* (Eye on Evaluation, E&R Report No. 03.24). Raleigh, NC: Wake County Public School System.
Participants	During the 2002/03 school year, 80 third-grade students received the *Fast ForWord®* program.[1] Of these, 71 were matched with students from non-*Fast ForWord®* schools based on race, limited English proficiency status, special programs code, free and reduced-price lunch status, and reading pretest scores. The remaining nine were missing either pre- or posttest scores. *Fast ForWord®* was used in six elementary schools, and the comparison students were taken from schools that did not use *Fast ForWord®*.
Setting	The students attended Wake County Public Schools in North Carolina.
Intervention	For the entire range of grades and students in the study, 91% used *Fast ForWord® Language* (first stage), 56% used *Fast ForWord® Language to Reading* (second stage), and 13% used *Fast ForWord® to Reading* (third stage); however, no information is provided by grade level.
Comparison	No information is provided.
Primary outcomes and measurement	North Carolina's End of Grade test was used as both the pretest and the posttest.
Teacher training	No information is provided.

1. The intervention was also used with seven students in grade 1 and 78 student in grade 2, but data were not presented for these students. The study also presented data for students in grades 4–8, attending a total of six elementary and four middle schools, but these students do not fall within the age range of the WWC's Beginning Reading review.

Appendix A2.1 Outcome measures in the alphabetics domain by construct

Outcome measure	Description
Phonological awareness	
Dynamic Indicators of Basic Early Literacy (DIBELS): Initial Sound Fluency subtest	This standardized test measures a child's ability to identify the initial sound in an orally presented word. The child is presented with four pictures and associated names and asked to identify the picture that starts with the same sound presented orally by the examiner (as cited in Scientific Learning Corporation, 2006).
Early Elementary version of the Test of Phonological Awareness (TOPA): Phonological Awareness subtest	The TOPA is a standardized, group-administered test designed to measure children's skill in identifying individual phonemes. The 10 ending sound-same items require children to identify which of three words ends with the same sound as a target word, while the 10 ending sound-different items ask children to mark which of a group of four words ends in a different sound than the others (as cited in Scientific Learning Corporation, 2005a).
Reading Edge: Initial Sound Discrimination subtest	This subtest of the Reading Edge test is a software-based assessment in a game format called "Jules Rampart Cooks with Gusto." It measures students' skills in segmenting words into phonemes or sounds units and recognizing and discriminating individual phonemes in common spoken words (as cited in Scientific Learning Corporation, 2006).[1]
Letter knowledge	
DIBELS: Letter Naming Fluency subtest	This is a subtest of a standardized measure in which students are presented with a page of upper- and lower-case letters arranged in a random order and are asked to name as many letters as they can. The score is the number of letters named correctly in one minute (as cited in Scientific Learning Corporation, 2006).
Phonics	
TOPA: Letter Sounds subtest	The TOPA is a standardized, group-administered test designed to measure children's skill in identifying the sounds of individual letters. The letter-sounds test requires children to spell simple pseudowords that are given as the names of "funny animals." The words vary from two to five phonemes in length, and they are all single-syllable. The child's score is the total number of words spelled correctly (as cited in Scientific Learning Corporation, 2005a).
Reading Edge: Initial Sound Knowledge subtest	This subtest of the Reading Edge test is a software-based assessment in a game format called "Squid Sisters." The initial sounds knowledge subtest measures children's skill in identifying the letter on the computer that corresponds to an orally-presented sound (as cited in Scientific Learning Corporation, 2006).[1]
Reading Edge: Non-Word Recognition subtest	This subtest of the Reading Edge test is a software-based assessment in a game format called "Squid Sisters." The non-word recognition subtest measures children's skill in decoding non-words by asking them to choose a correct word from a group of other nonsense words. The words start with one syllable and increase in difficulty (as cited in Scientific Learning Corporation, 2006).[1]
Test of Word Reading Efficiency (TOWRE): Phonemic Decoding Efficiency subtest	The TOWRE is a standardized, nationally normed measure. The phonetic decoding efficiency subtest measures the number of pronounceable printed non-words that can be accurately decoded within 45 seconds (as cited in Scientific Learning Corporation, 2005b).
TOWRE: Sight Word Efficiency subtest	The TOWRE is a standardized, nationally normed measure. The sight word efficiency subtest assesses the number of real printed words that can be accurately identified within 45 seconds (as cited in Scientific Learning Corporation, 2005b).
Woodcock Johnson (WJ): Letter Word Identification subtest	The WJ letter word identification subtest measures a student's skill in identifying individual letters and words (as cited in Scientific Learning Corporation, 2006).

(continued)

Appendix A2.1 Outcome measures in the alphabetics domain by construct *(continued)*

Outcome measure	Description
Phonological awareness and phonics	
TOPA	The TOPA measures children's skills in identifying individual phonemes at the beginning and end of orally-presented words and in writing letters when given letter sounds (as cited in Scientific Learning Corporation, 2006).

1. The Reading Edge test was developed by Scientific Learning Corporation, which also developed *Fast ForWord®*. While there is no evidence of obvious overalignment between the measure and the intervention (intervention student receiving exposure to the measure during the course of treatment), it should be noted that the developer of the intervention and the measure were the same.

Appendix A2.2 Outcome measures in the comprehension domain

Outcome measure	Description
Reading comprehension	
Comprehensive Test of Basic Skills, Fifth Edition (CTBS/5) Terra Nova: Total Reading subtest	A group-administered, standardized assessment of reading comprehension (as cited in Borman & Benson, 2006).
Degrees of Reading Power (DRP)	An untimed, standardized test requiring students to read a non-fiction passage with a word or set of words missing. Students select an appropriate answer to complete the sentence from a set of four or five alternatives (as cited in Scientific Learning Corporation, 2005c).
North Carolina End of Grade Test	A standardized state assessment designed to match the North Carolina curriculum. It uses multiple-choice questions with reading passages and is designed to measure comprehension skills. Students read eight reading selections of varying genres and answer three to nine comprehension questions for each (as cited in Overbay & Baenen, 2003).

Appendix A3.1 Summary of findings for the alphabetics domain by construct[1]

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome (standard deviation[2]): *Fast ForWord®* group	Authors' findings from the study: Mean outcome (standard deviation[2]): Comparison group	WWC calculations: Mean difference[3] (*Fast ForWord®* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Construct: Phonological awareness								
Scientific Learning Corporation, 2005a (randomized controlled trial)[7]								
Early Elementary version of TOPA: Phonological Awareness subtest	First and second grade	197	53.7 (25.0)	46.8 (25.7)	6.9	0.27	Statistically significant	+11
Scientific Learning Corporation, 2006 (randomized controlled trial)[7]								
DIBELS: Initial Sound Fluency subtest	Kindergarten	47	15.1 (8.5)	20.0 (9.8)	–4.9	–0.53	ns	–20
Reading Edge: Initial Sound Discrimination subtest	Kindergarten	43	29.4 (17.6)	23.4 (13.2)	6.0	0.38	ns	+15
Construct: Letter knowledge								
Scientific Learning Corporation, 2006 (randomized controlled trial)[7]								
DIBELS: Letter Naming Fluency subtest	Kindergarten	47	26.5 (11.5)	28.2 (11.0)	–1.7	–0.15	ns	–6
Construct: Phonics								
Scientific Learning Corporation, 2005a (randomized controlled trial)[7]								
Early Elementary version of TOPA: Letter Sounds subtest	First and second grade	197	42.7 (18.4)	38.9 (19.3)	3.8	0.20	Statistically significant	+8
Scientific Learning Corporation, 2005b (randomized controlled trial)[7]								
TOWRE: Phonemic Decoding Efficiency subtest	Third grade	50	107.9 (15.1)	103.1 (12.3)	4.7	0.34	ns	+13
TOWRE: Sight Word Efficiency subtest	Third grade	50	99.9 (16.1)	96.3 (10.4)	3.6	0.26	ns	+10
Scientific Learning Corporation, 2006 (randomized controlled trial)[7]								
Reading Edge: Initial Sound Knowledge subtest	Kindergarten	43	61.5 (44.1)	58.8 (42.2)	2.7	0.06	ns	+2

(continued)

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

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome (standard deviation[2]) *Fast ForWord®* group	Comparison group	WWC calculations: Mean difference[3] (*Fast ForWord®* – comparison)	Effect size[4]	Statistical significance[5] (at $\alpha = 0.05$)	Improvement index[6]
Reading Edge: Non-Word Recognition subtest	Kindergarten	41	15.4 (15.5)	12.5 (14.5)	2.9	0.19	ns	+8
Woodcock Johnson: Letter Word Identification subtest	Kindergarten	48	106.1 (9.9)	103.7 (7.4)	2.4	0.27	ns	+11
Construct: Phonological awareness and phonics								
Scientific Learning Corporation, 2006 (randomized controlled trial)[7]								
TOPA	Kindergarten	47	106.0 (11.7)	105.0 (11.7)	1.0	0.08	ns	+3
Averages[8]								
Average for alphabetics (Scientific Learning Corporation, 2005a)						0.24	Statistically significant	+9
Average for alphabetics (Scientific Learning Corporation, 2005b)						0.30	Statistically significant	+12
Average for alphabetics (Scientific Learning Corporation, 2006)						0.04	ns	+2
Domain average for alphabetics						0.19	na	+8

ns=not statistically significant
na=not applicable
TOPA = Early Elementary version of the Test of Phonological Awareness
DIBELS = Dynamic Indicators of Basic Early Literacy
TOWRE = Test of Word Reading Efficiency

1. Appendix A3.1 reports overall findings considered for the rating of effectiveness and the average 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. Pre- and posttest standard deviations were requested by the WWC for the Scientific Learning Corporation studies for purpose of effect size calculation and were received from the study author.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The means and mean difference are regression-adjusted to control for differences in the pretest in Scientific Learning Corporation 2005a, 2005b, and 2006, using data requested by the WWC and provided by the author.
4. For an explanation of effect size calculation, please see the 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 the Scientific Learning Corporation 2005a, 2005b, and 2006 studies, no corrections for clustering were needed. The effect sizes for these three studies were based on data provided to the WWC by the authors. For Scientific Learning Corporation (2005a), the authors reported joint significance for the two TOPA subtests; subsequent author calculations reported directly to the WWC showed individual significance, which was verified by the WWC after correcting for multiple comparisons. For Scientific Learning Corporation (2005b), the authors reported joint significance for the two TOWRE subtests; however, subsequent author calculations reported directly to the WWC showed that the individual subtests were not statistically significant, so no corrections for multiple comparisons were made.
8. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The improvement index for the domain is calculated from the domain average effect size.

Appendix A3.2 Summary of findings for the comprehension domain[1]

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome (standard deviation[2]): *Fast ForWord®* group	Authors' findings from the study: Mean outcome (standard deviation[2]): Comparison group	WWC calculations: Mean difference[3] (*Fast ForWord®* – comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Construct: Reading comprehension								
Borman & Benson, 2006 (randomized controlled trial)[7]								
Comprehensive Test of Basic Skills, Fifth Edition, Terra Nova: Total Reading subtest	Second grade	112	33.4 (15.9)	34.6 (12.4)	–1.2	–0.14	ns	–6
Scientific Learning Corporation, 2005c (randomized controlled trial)[7]								
Degrees of Reading Power	Second and third grade	38	41.9 (15.8)	33.2 (18.8)	8.8	0.51	Statistically significant	+19
Overbay & Baenen, 2003 (quasi-experimental design)[7]								
North Carolina End of Grade Test	Third grade	142	243.2 (np)[8]	245.9 (np)[8]	–2.8	–0.32	ns	–12
Domain average for comprehension[9]						0.03	na	+1

ns=not statistically significant
na=not applicable
np = not presented

1. Appendix A3.2 reports overall findings considered for the rating of effectiveness and the average 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. Pre- and posttest standard deviations were requested by the WWC for the Scientific Learning Corporation study for the purpose of effect size calculation and were received from the study author.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. The means and mean difference are regression-adjusted to control for differences in the pretest in Scientific Learning Corporation (2005c), using data provided by the author in a personal communication.
4. For an explanation of effect size calculation, please see the Technical Details of WWC-Conducted Computations. The effect size for Borman & Benson (2006) was provided by the authors using an alternative, though acceptable, method of calculation; the authors used a difference in gain score divided by the pooled posttest standard deviation.
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 Borman & Benson (2006), Scientific Learning Corporation (2005c), and Overbay & Baenen (2003), no corrections for clustering or multiple comparisons were needed. The effect size for Scientific Learning Corporation (2005c) was based on data provided to the WWC by the authors.
8. The standard deviation was not presented in the study.
9. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The improvement index for the domain is calculated from the domain average effect size.

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

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome (standard deviation): *Fast ForWord®* group	Authors' findings from the study: Mean outcome (standard deviation): Comparison group	WWC calculations: Mean difference[2] (*Fast ForWord®* – comparison)	WWC calculations: Effect size[3]	WWC calculations: Statistical significance[4] (at $\alpha = 0.05$)	WWC calculations: Improvement index[5]
Construct: Phonological awareness								
Scientific Learning Corporation, 2005d (randomized controlled trial)[6]								
Early Elementary version of the Test of Phonological Awareness (TOPA): Letter Sounds subtest	Springfield, Ohio: First and second grade	93	43.1 (16.4)	36.6 (18.31)	6.2	0.36	Statistically significant	+14
TOPA: Phonological Awareness subtest	Springfield, Ohio: First and second grade	93	57.9 (24.6)	44.8 (25.8)	8.6	0.34	Statistically significant	+13

1. This appendix presents subgroup findings for one of three districts presented in Scientific Learning Corporation (2005a) for measures in the alphabetics domain. Total group scores were used for rating purposes and are presented in Appendix A3.1.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
3. For an explanation of effect size calculation, please see the Technical Details of WWC-Conducted Computations.
4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
5. 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.
6. 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 Scientific Learning Corporation (2005d), no corrections for clustering were needed.

Appendix A5.1 *Fast ForWord*® 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 *Fast ForWord*® as having positive effects. It met the criteria for positive effects because two studies with strong designs showed statistically significant positive effects, and only one study showed an indeterminate effect. The remaining ratings (potentially positive effects, mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *Fast ForWord*® was assigned the highest possible rating.

Rating received
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.
Met. Two studies of *Fast ForWord*® had a statistically significant positive effect in this domain, and both met WWC evidence standards for a strong design.
AND
• Criterion 2: No studies showing statistically significant or substantively important *negative* effects.
Met. No studies of *Fast ForWord*® showed statistically significant or substantively important negative 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.

Appendix A5.2 *Fast ForWord®* 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 *Fast ForWord®* as having mixed effects. It did not meet the criteria for positive effects because no studies showed statistically significant positive effects. In addition, it did not meet the criteria for potentially positive effects because two studies showed indeterminate effects and only one study showed a substantively important positive effect. The remaining ratings (no discernible effects, potentially negative effects, and negative effects) were not considered because *Fast ForWord®* was assigned the highest applicable rating.

Rating received

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.

  Met. One study had substantively important and positive effects and one study had substantively important and negative 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. One study had substantively important and positive effects, one study had substantively important and negative effects, and one study 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 *Fast ForWord®* showed statistically significant positive effects in this domain.

AND

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

  Met. No studies of *Fast ForWord®* showed 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.

  Met. One study of *Fast ForWord®* had a substantively important positive effect in this domain.

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.

  Not met. One study had substantively important and 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 A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence[1]
		Schools	Students	
Alphabetics	3	5	295	Small
Fluency	0	0	0	na
Comprehension	3	>11	292	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."