# What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Beginning Reading January 2009[1]

# Earobics®

## Program description[2]

*Earobics*® is interactive software that provides students in pre-K through third grade with individual, systematic instruction in early literacy skills as students interact with animated characters. *Earobics*® *Foundations* is the version for prekindergaten, kindergarten, and first grade. *Earobics*® *Connections* is for second and third graders and older struggling readers. The program builds children's skills in phonemic awareness, auditory processing, and phonics, as well as the cognitive and language skills required for comprehension. Each level of instruction addresses recognizing and blending sounds, rhyming, and discriminating phonemes within words, adjusting to each student's ability level. The software is supported by music, audiocassettes, and videotapes and includes picture/word cards, letter-sound decks, big books, little books, and leveled readers for reading independently or in groups.

## Research

Two studies of *Earobics*® meet What Works Clearinghouse (WWC) evidence standards and two studies meet WWC evidence standards with reservations. The four studies included 246 students from grades K through 3 in Los Angeles, California; southwest Florida; Anchorage, Alaska; and Chicago, Illinois.[3]

Based on these four studies, the WWC considers the extent of evidence for *Earobics*® to be small for alphabetics and reading fluency. No studies that meet WWC evidence standards with or without reservations examined the effectiveness of *Earobics*® in the comprehension or general reading achievement domains.

## Effectiveness

*Earobics*® was found to have positive effects on alphabetics and potentially positive effects on reading fluency.

| | *Alphabetics* | *Reading fluency* | *Comprehension* | *General reading achievement* |
|---|---|---|---|---|
| **Rating of effectiveness** | Positive effects | Potentially positive effects | na | na |
| **Improvement index[4]** | Average: +25 percentile points<br>Range: 0 to +49 percentile points | Average: +15 percentile points<br>Range: +3 to +33 percentile points | na<br>na | na<br>na |

na = not applicable

1. This report has been updated to include reviews of 10 studies released in 2005 and later. A complete list and disposition of all studies reviewed is provided in the references.
2. The descriptive information for this program was obtained in 2007 from a publicly available source: the program's website (http://www.earobics.com). The WWC requests developers to review the program description sections for accuracy from their perspective. Neither the authors nor the website provided any additional information for this update. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review.
3. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.
4. These numbers show the average and range of student-level improvement indices for all findings across the studies.

## Additional program information

### Developer and contact

Developed in 1995, *Earobics*® is distributed by Houghton Mifflin Harcourt Learning Technology. Address: Earobics | Houghton Mifflin Harcourt Learning Technology, 222 Berkeley Street, Boston, MA 02116. Email: HMLTcustomerservice@hmco.com. Web: www.earobics.com. Telephone: (888) 242-6747.

### Scope of use

According to the distributors, *Earobics*® has been used nationally in more than 10,000 schools. The program has been used with at-risk students, general and special education students, and English language learners.

### Teaching

The software is a supplemental program that can be used in conjunction with existing language arts programs. The *Earobics*® Teacher's Guides help teachers plan students' use of the software and supporting materials, using a teach, practice, and apply approach. As students work with the software, the program automatically adjusts based on each student's performance. Reports on student performance can be printed or accessed online. Teachers may also customize the program for students, including selecting one of 10 languages for the directions. Teachers also have access to CD-ROMs with reproducible materials tied to specific lessons for students. Professional development for using *Earobics*® is available and focuses on instructional strategies to incorporate *Earobics*® into the curricula.

### Cost[5]

*Earobics*® *Foundations* and *Earobics*® *Connections* are available for home use for $59 per user or in a "clinic" version that accommodates up to 12 users for $299. *Foundations* is targeted for ages 4 to 7 and includes six interactive games with more than 300 levels of play. *Connections* is targeted for ages 7 to 10 and includes five interactive games with nearly 600 levels of play.

## Research

Twenty-eight studies reviewed by the WWC investigated the effects of *Earobics*®. Two studies (Cognitive Concepts, 2003, and Gale, 2006) are randomized controlled trials that meet WWC evidence standards. Two studies (Rehmann, 2005, and Valliath, 2002) are randomized controlled trials or quasi-experimental designs that meet WWC evidence standards with reservations. The remaining 24 studies do not meet either WWC evidence standards or eligibility screens.

### Meets evidence standards

Cognitive Concepts (2003) conducted a randomized controlled trial of elementary school students in Los Angeles, California. Nineteen teachers identified students in kindergarten through third grade with reading difficulties. Students were pretested, matched, and then randomly divided into two groups. In all, 39 students used *Earobics*® in addition to *Open Court* (their regular reading curriculum) during the study period that lasted from October to December, and 35 students in the comparison group used only *Open Court*.

Gale (2006) identified kindergarten and first-grade students whose fall Dynamic Indicators of Basic Early Literacy Skills (DIBELS) test scores indicated that they needed substantial intensive intervention. Among those students, 41 kindergarten students and 38 first-grade students eligible to participate in the study returned parental consent forms. These students were randomly assigned to one of three groups: (1) *Earobics*® Step 1, (2) Lexia Early Reading or (3) control. Students in groups 1 and 2 received the supplemental interventions during the five week study period in addition to their regular instruction, while the control group received no instruction beyond their regular language arts class time.

### Meets evidence standards with reservations

Rehmann (2005) was a randomized controlled trial with severe differential attrition. At a Title I school with 140 kindergarten

5. Distributor's prices as of August 2007.

## Research *(continued)*

and first-grade students (70 at each grade level), the researcher blocked the students by gender and grade, randomly selected 20 students from each block to participate in the study, and then randomly assigned the 80 students from all four blocks to receive the 10-week intervention (40 students) or to a comparison group (40 students).[6] Among this sample, 14 were discontinued during the study (10 in the intervention group and 4 in the comparison group), leaving a final analysis sample of 66 students.

Valliath (2002) conducted a quasi-experimental study of first-grade students from three elementary public schools in a high-achieving school district in Chicago, Illinois. Ten teachers identified three children with the lowest reading ability within their respective classrooms. Students were pretested, matched, and divided into two similar groups. In the analysis sample, 15 students used six exercises of the *Earobics*® software for ten weeks, while 15 students in the comparison group used math software.

### Extent of evidence

The WWC categorizes the extent of evidence in each domain as small or medium 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 meet WWC evidence standards with or without reservations.[7]

The WWC considers the extent of evidence for *Earobics*® to be small for alphabetics and reading fluency. No studies that meet WWC evidence standards with or without reservations examined the effectiveness of *Earobics*® in the comprehension or general reading achievement domains.

## Effectiveness

### Findings

The WWC review of interventions for Beginning Reading addresses student outcomes in four domains: alphabetics, reading fluency, comprehension, and general reading achievement. The studies included in this report cover two domains: alphabetics and reading fluency. The findings below present the authors' estimates and WWC-calculated estimates of the size and the statistical significance of the effects of *Earobics*® on students.[8]

*Alphabetics.* Four studies presented findings in the alphabetics domain. Cognitive Concepts (2003) found statistically significant positive effects on three phonological awareness measures (ORAL-J: Blending into Words, Segmenting into Sounds, and Rhyming Words subtests).[9] The study authors did not find statistically significant effects of *Earobics*® on the letter knowledge measure (ORAL-J: Letter Naming subtest) or the phonics measure (ORAL-J: Sound of Letters subtest). The average effect size across the five outcomes was large enough to be considered substantively important according to WWC criteria (that is, an effect size of at least 0.25).

Gale (2006) analyzed three alphabetics outcomes (DIBELS: Initial Sounds Fluency, Letter Naming Fluency, and Phoneme Segmentation Fluency subtests) for kindergarten students and

6. Students in the control group received the intervention in a second phase of the study, and the students in the original intervention group served as the comparison for that phase. The WWC focuses on the first phase, because by the second phase, the comparison group had just received the intervention.
7. 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. Information about how the extent of evidence rating was determined for *Earobics*® is in Appendix A5.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate the statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Cognitive Concepts (2003), Rehmann (2005), and Valliath (2002), no corrections for clustering or multiple comparisons were needed. In the case of Gale (2006), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.
9. Data for these three phonological awareness measures were received through communication with the study author.

## Effectiveness *(continued)*

three outcomes (DIBELS: Letter Naming Fluency, Phoneme Segmentation Fluency, and Nonsense Word Fluency subtests) for first-grade students. In examining *Earobics*® versus the group that received no supplemental instruction, the author found, and the WWC confirmed, positive and statistically significant effects of *Earobics*® for three DIBELS subtests: Initial Sounds Fluency (kindergarten), Phoneme Segmentation Fluency (kindergarten and grade 1) and Nonsense Words Fluency (kindergarten and grade 1). There were no statistically significant effects on the DIBELS: Letter Naming Fluency subtest in kindergarten or grade 1. However, the WWC determined that the effects for both grades were large enough to be considered substantively important. The study also compared *Earobics*® to Lexia, and the author found no statistically significant effect on any of the four DIBELS subtests for either of the two grades. However, the WWC determined that three of the positive effects were large enough to be considered substantively important: Initial Sounds Fluency (kindergarten), Phoneme Segmentation Fluency (grade 1) and Nonsense Words Fluency (grade 1). The WWC found that the combined effect for alphabetics across both comparison groups was not statistically significant. However, the WWC determined that the combined effect was large enough to be considered substantively important.

Rehmann (2005) found that the overall *Earobics*® effect across the three of four alphabetics measures of beginning reading (DIBELS: Initial Sound Fluency, Letter Naming Fluency, and Phoneme Segmentation Fluency) was not statistically significant. For one subtest (Nonsense Words Fluency), the WWC determined that the negative effect was substantively important (that is, an effect size with an absolute value of at least 0.25).

Valliath (2002) found that the overall intervention effect across the eight measures of beginning reading was not statistically significant.[10] The WWC analyzed four phonological awareness measures (Comprehensive Test of Phonological Processing [CTOPP]: Blending Words, Blending Non-Words, Elision, and Sound Matching subtests) and two phonics measures (Woodcock Reading Mastery Test: Word Identification and Word Attack subtests). The WWC found that the effect for one of the four phonological awareness tests (CTOPP: Sound Matching subtest) was positive and statistically significant. Effects for the other three phonological awareness subtests and the two phonics subtests were not statistically significant but the WWC determined them to be substantively important. The average effect size across the six outcomes was large enough to be considered substantively important according to the WWC criteria.

*Reading fluency.* Two studies presented findings in the reading fluency domain. In analyzing ORAL-J: Words per Minute subtest data, Cognitive Concepts (2003) did not find statistically significant effects of *Earobics*®, and the effect was not large enough to be considered substantively important according to WWC criteria. Gale (2006) found positive but not statistically significant effects of *Earobics*® when compared to Lexia and the no intervention group on the DIBELS: Oral Reading Fluency subtest. The WWC determined that both positive effects were large enough to be 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, 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).

10. The WWC did not use all eight measures in its analysis because two were outside the domain specified in the beginning reading protocol. See Appendix A1.4.

**The WWC found *Earobics*® to have positive effects for alphabetics and potentially positive effects for reading fluency**

### 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 +25 percentile points across the four studies, with a range of 0 to +49 percentile points across findings.

The average improvement index for reading fluency is +15 percentile points across two studies, with a range of +3 to +33 percentile points across findings.

### Summary

The WWC reviewed 28 studies on *Earobics*®. Two of these studies meet WWC evidence standards; two studies meet WWC evidence standards with reservations; the remaining 24 studies either do not meet WWC evidence standards or do not meet eligibility screens. Based on the four studies, the WWC found positive effects for alphabetics and potentially positive effects for reading fluency. The conclusions presented in this report may change as new research emerges.

## References


### Meets WWC evidence standards

Cognitive Concepts, Inc. (2003). *Outcomes report: Los Angeles Unified School District, California*. Retrieved from http://www.earobics.com/results/la.php.

Gale, D. (2006). *The effect of computer-delivered phonological awareness training on the early literacy skills of students identified as at-risk for reading failure.* Retrieved from the University of South Florida website: http://purl.fcla.edu/usf/dc/et/SFE0001531.

### Meets WWC evidence standards with reservations

Rehmann, R. (2005). *The effect of Earobics (TM) Step 1, software on student acquisition of phonological awareness skills.* Dissertation. *Dissertation Abstracts International, 66*(07A), 157–2533. (UMI No. 3181124)

Valliath, S. (2002). *An evaluation of a computer-based phonological awareness training program: Effects on phonological awareness, reading and spelling. Dissertation Abstracts International, 63*(04), 1291A. (UMI No. 3050601)

### Studies that fall outside the Beginning Reading protocol or do not meet WWC evidence standards

Cognitive Concepts, Inc. (2000). *Earobics Early Literacy Instruction: Chicago Public Schools pilot research report.* Retrieved from http://www.earobics.com/results/chicago.php. This study is ineligible for review because it does not use a comparison group.

Cognitive Concepts, Inc. (2001). *Outcomes report: Daviess County Public Schools, Kentucky*. Retrieved from http://www.earobics.com/results/additional_reports.php. This study is ineligible for review because it does not use a comparison group.

Cognitive Concepts, Inc. (2001). *Outcomes report: Newport News Public Schools, Virginia.* Retrieved from http://www.earobics.com/results/newport.php. This study is ineligible for review because it does not use a sample within the age or grade range specified in the protocol.

Cognitive Concepts, Inc. (2001). *Outcomes report: PALS assessment, Virginia.* Retrieved from http://www.earobics.com/

## References *(continued)*


results/newportpals.php. This study is ineligible for review because it does not use a comparison group.

Cognitive Concepts, Inc. (2001). *Outcomes report: Spring Branch Independent School District, Texas*. Retrieved from http://www.earobics.com/results/spring.php. This study is ineligible for review because it does not use a comparison group.

Cognitive Concepts, Inc. (2002). *Outcomes report: Anne Arundel County Public Schools, Maryland*. Retrieved from http://www.earobics.com/results/additional_reports.php. This study is ineligible for review because it does not use a comparison group.

Cognitive Concepts, Inc. (2002). *Outcomes report: Brevard County Public Schools, Florida*. Retrieved from http://www.earobics.com/results/additional_reports.php. This study is ineligible for review because it does not use a comparison group.

Cognitive Concepts, Inc. (2002). *Outcomes report: Cincinnati Children's Hospital Medical Center, Ohio*. Retrieved from http://www.earobics.com/results/additional_reports.php. This study is ineligible for review because it does not use a comparison group.

Cognitive Concepts, Inc. (2002). *Outcomes report: Culver City Unified School District, California.* Retrieved from http://www.earobics.com/results/additional_reports.php. This study is ineligible for review because it does not use a comparison group.

Cognitive Concepts, Inc. (2002). *Outcomes report: District of Columbia Public Schools, Washington, DC.* Retrieved from http://www.earobics.com/results/columbia.php. This study is ineligible for review because it does not use a comparison group.

Cognitive Concepts, Inc. (2002). *Outcomes report: Northwestern University, Illinois.* Retrieved from http://www.earobics.com/results/additional_reports.php. This study is ineligible for review because it does not use a comparison group.

Cognitive Concepts, Inc. (2002). *Outcomes report: Polk County School District, Florida.* Retrieved from http://www.earobics.com/results/polk.php. This study is ineligible for review because it does not use a comparison group.

Hargett, K. (2005). *The effects of computer-assisted therapy on phonological awareness in school-aged children*. Unpublished manuscript, Louisiana Tech University. The study is ineligible for review because it does not use a comparison group.

Hayes, E. A., Warrier, C. M., Nicol, T. G., Zecker, S. G., & Kraus, N. (2002). Neural plasticity following auditory training in children with learning problems. *Clinical Neurophysiology, 114,* 673–684. This study is ineligible for review because it does not include a student outcome.

Houghton Mifflin *Earobics. New Orleans education closes achievement gap with Earobics.* Retrieved March 3, 2008, from www.earobics.com. The study is ineligible for review because it does not use a comparison group.

Kamhi, A. G. (2006). Epilogue: Some final thoughts on EBP. *Language, Speech, and Hearing Services in Schools, 37*(4), 320–322. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Moore, D. R., & Amitay, S. (2007). Auditory training: Rules and applications. *Seminars in Hearing, 28,* 99–109. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention.

Pettis, A. M. (2000). *A study on phonological awareness: The comparison of two computer-based programs used as intervention for students with disabilities.* Unpublished master's thesis, Grand Valley State University, Allendale, MI. The study does not meet evidence standards because the overall attrition rate exceeds WWC standards for this area.

Pobanz, M. S. (2000). *The effectiveness of an early literacy/auditory processing training program, called Earobics, with young children achieving poorly in reading*. Paper presented at the meeting of the California Association of Social Psychologists, Los Angeles, CA. This study is ineligible for review because it does not use a comparison group.

Pokorni, J. L., Worthington, C. K., & Jamison, P. J. (2004). Phonological awareness intervention: Comparison of *Fast ForWord,*

## References *(continued)*

*Earobics,* and *LiPS. The Journal of Educational Research, 97*(3), 147–157. This study is ineligible for review because it does not disaggregate findings for the age or grade range specified in the protocol.

Purpura, D. J. (2007). Can externalizing behaviors be altered by an early reading intervention. Retrieved from http://etd.lib.fsu.edu/theses/available/etd-07092007-145531/unrestricted/Purpura_thesis.pdf. The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.

Rack, C. J. (2006). *The effects of using Earobics and learning resources phonics based software as supplemental computer assisted instruction on struggling second grade readers' phonemic awareness, phonics skills, and word recognition.* Unpublished manuscript. Cardinal Stritch University: Milwaukee, WI. The study does not meet WWC evidence standards because the measures of effect cannot be attributed solely to the intervention—the intervention was combined with another intervention.

Richard, G. J. (2005). Language-based assessment and intervention of APD. In Teralandur K. Parthasarathy (Ed.), *An introduction to auditory processing disorders in children* (pp. 95–108). Hillsdale, NJ: Erlbaum. The study is ineligible for review because it does not examine the effectiveness of an intervention.

Sweeney, D. P., & Hoffman, C. D. (2007). Research issues in autism spectrum disorders. In Robert B. Rutherford Jr., Mary Magee Quinn, and Sarup R. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 308–322). New York, NY: Guilford. The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.

# Appendix

## Appendix A1.1 Study Characteristics: Cognitive Concepts, 2003 (randomized controlled trial)

| Characteristic | Description |
|---|---|
| **Study citation** | Cognitive Concepts, Inc. (2003). *Outcomes report: Los Angeles Unified School District, California.* Retrieved from http://www.earobics.com/results/la.php. |
| **Participants** | Nineteen teachers identified students in kindergarten through third grade with reading difficulties. More than 80% of students were English language learners. The study author administered pretests (ORAL-J and Test of Memory and Learning [TOMAL]) to students to divide them into two similar groups.[1] The groups were then randomly assigned to be either the intervention or comparison groups. Each group originally had 43 students, but there was some attrition due to poor attendance.[2] In the analysis sample, 39 students were in the treatment group and 35 students were in the comparison group. |
| **Setting** | The study took place in one elementary school located Los Angeles, California. |
| **Intervention** | Students in the intervention group were given directions on how to use *Earobics®* software. They received instruction with *Earobics®* for 30 minutes a day, five days a week from October through December. In addition, the intervention group received its regular whole class reading instruction with the *Open Court Reading* curriculum. |
| **Comparison** | Students in the comparison classes received the regular whole class reading instruction with the *Open Court Reading* curriculum during the language arts period. |
| **Primary outcomes and measurement** | For both pre- and posttests, the authors administered eight subtests of the ORAL-J: Early Literacy Achievement test: Blending into Words, Segmenting into Sounds, Rhyming Words, Letter Naming, and Sound of Letters subtests, as well as three administrations of the Words per Minute subtest.[3] The TOMAL was also used in the study, but it was not included in this review because it was outside the scope of the Beginning Reading review. For a more detailed description of these outcome measures, see Appendices A2.1 and A2.2. |
| **Staff/teacher training** | No information on teacher training is provided. The *Earobics®* group worked in a computer lab, with minimal teacher instruction. |

1. Equivalence of the two groups at pretest was confirmed through data sent by the author, M. Poblanz.
2. Some information about attrition was provided through personal communication with the author.
3. Some of the test data were not in the published report and were provided directly by the author.

## Appendix A1.2 Study Characteristics: Gale, 2006 (randomized controlled trial)

| Characteristic | Description |
|---|---|
| **Study citation** | Gale, D. (2006). *The effect of computer-delivered phonological awareness training on the early literacy skills of students identified as at-risk for reading failure.* Retrieved from the University of South Florida website: http://purl.fcla.edu/usf/dc/et/SFE0001531. |
| **Participants** | Kindergarten and first-grade students who were identified in the fall assessment period as needing intensive substantial intervention based on their performance on the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) were recruited for this study. Forty-one kindergarten students and 38 first-grade students eligible to participate returned parental consent forms. These students were randomly assigned to one of three groups: (1) *Earobics®* Step 1, (2) Lexia Early Reading, or (3) control. After attrition, the analysis sample contained 39 kindergarten and 37 first-grade students. |
| **Setting** | The elementary school in which this study occurred is located in a large school district in the southwest region of Florida serving approximately 114,466 pre-K to 12th-grade students. The elementary school has a total kindergarten through fifth-grade student enrollment of 722. Students in the school represent the following ethnic groups: 60% Caucasian, 19% Hispanic, 8% Asian/Pacific Islander, 7% African American, 5% multiracial, <1% American Indian/Alaskan Native. Approximately 73% of the students in this school are eligible for free or reduced-priced lunch. |
| **Intervention** | A rotation schedule was developed by the researcher based on teacher input. The two phonological awareness software programs were loaded on 14 numbered computers with headphones in the computer lab at the elementary school. Each student was assigned to a computer to use throughout the intervention period. Before the intervention period began, the researcher trained the participants in small groups of five on the relevant intervention software (*Earobics®* Step 1 or Lexia Early Reading) with regard to initiating and proceeding through the program and navigating the mouse. Students were required to pass at least five out of six areas on the training checklist as well as the task "use mouse to navigate activity" before beginning the intervention. The students were divided into four groups that alternated in the computer lab according to the rotation schedule. The researcher and a teacher assistant monitored the students each day during their training in the computer lab. Students used their respective computer programs in the school computer lab 20 minutes daily for 25 days, resulting in a total of 8 hours 20 minutes of exposure. |
| **Comparison** | The control group received no specific intervention designated by the study. Typical reading instruction in the school was a 90-minute reading block. |
| **Primary outcomes and measurement** | Students were tested before and after the intervention using the DIBELS subtests for Initial Sounds Fluency (kindergarten only), Letter Naming Fluency, Phoneme Segmentation Fluency, Nonsense Word Fluency (first grade only) and Oral Reading Fluency (first grade only). For a more detailed description of these outcome measures, see Appendices A2.1 and A2.2. |
| **Staff/teacher training** | No information on teacher training is provided. The *Earobics®* group worked in a computer lab, with minimal teacher instruction. |

## Appendix A1.3 Study Characteristics: Rehmann, 2005 (randomized controlled trial)

| Characteristic | Description |
|---|---|
| **Study citation** | Rehmann, R. (2005). *The effect of Earobics (TM) Step 1, software on student acquisition of phonological awareness skills.* Dissertation. *Dissertation Abstracts International, 66*(07A), 157–2533. (UMI No. 3181124) |
| **Participants** | At a school with 140 kindergarten and first-grade students (70 at each grade level), the researcher blocked the students by gender and grade, and then randomly selected a sample of 80 students (40 in group 1, 40 in group 2). Among this sample, 14 were discontinued during the study (10 in group 1 and four in group 2), leaving a final analysis sample of 66 students. Students in the study participated in two 10-week intervention phases. During the first phase, group 1 received the intervention and group 2 served as the comparison. In the second phase, group 2 received the intervention and group 1 was the comparison. (The WWC focuses on phase 1 only, because by phase 2, the comparison group had just received the intervention.) |
| **Setting** | One Title 1 elementary school in Anchorage, Alaska, participated in the study. The intervention was administered in the computer lab or in the student's regular classroom. |
| **Intervention** | The intervention group received computerized instruction in phonological awareness with *Earobics*® Step 1 software 20 minutes a day, three days a week, for a total of 10 weeks. *Earobics*® Step 1 uses a game format designed to assist students in developing specific phonological awareness and auditory-processing skills. The software consists of six multileveled interactive games with adaptive technology. This was in addition to whole-group direct instruction that the students received together with comparison students. |
| **Comparison** | While intervention students were engaged with the intervention software, comparison students received an additional 20 minutes of peer or individual classroom activities in a variety of formats dependent on the individual teacher's program. |
| **Primary outcomes and measurement** | Students were tested before and after the intervention using the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) 6th edition and the Comprehensive Test of Phonological Processing (CTOPP).[1] For a more detailed description of these outcome measures, see Appendices A2.1 and A2.2. |
| **Staff/teacher training** | Training was provided in two phases: implementation and data collection. Seven classroom teachers and one Title 1 paraprofessional received training, which was provided by specialists from the developer. Other staff members were trained and available to provide technical support, and additional support was available online and by telephone. |

1. The CTOPP was administered twice before the intervention and again after both groups 1 and 2 had been subject to the intervention, but not in the interim time when one group had the intervention and one did not. Therefore, there is no comparison group for this measure, and the results of the CTOPP for this study are not reported in Appendix A3.1. Additionally, not every DIBELS subtest was administered to each grade in the pretest (October) and posttest (December) period for the 10-week intervention phase. This intervention report describes only the subtests with scores at both administrations.

## Appendix A1.4 Study Characteristics: Valliath, 2002 (quasi-experimental design)

| Characteristic | Description |
|---|---|
| **Study citation** | Valliath, S. (2002). An evaluation of a computer-based phonological awareness training program: Effects on phonological awareness, reading and spelling. *Dissertation Abstracts International*, *63*(04), 1291A. (UMI No. 3050601) |
| **Participants** | Ten teachers identified three children with the lowest reading ability within their first-grade classrooms. Before pretesting, all 30 students received a score of at least 80 on the Test of Nonverbal Intelligence -3. Then the study author administered a pretest (the Woodcock Reading Mastery Test: Word Identification subtest) to students to divide them into two similar groups.[1] All children came from English-speaking monolingual homes; none received any special education or speech and language services. The sample consisted of 16 boys and 14 girls, ages 6.5 to 7.5. In the analysis sample, 15 students were in the intervention group and 15 students were in the comparison group. |
| **Setting** | The study took place in three public elementary schools from a high-achieving school district of a northwest suburb of Chicago, Illinois. |
| **Intervention** | Students in the intervention group spent 20 minutes each day playing one of the six *Earobics®* games. *Earobics®* delivered phonological awareness training in the auditory mode and provided minimal sound-to-print training. The children played individually and were provided headsets. They started at the lowest skill level for each game and progressed at their own pace. The games were rotated systematically on a daily basis during the 10-week training program. The average number of days attended by the students in the intervention group was 46.47 of a possible 50 days. |
| **Comparison** | Students in the comparison group received comparable amounts of daily exposure (approximately 20 minutes) to math training software, *Knowledge Adventure's Jump Start Math for First Graders.* The software has no linguistic training component and consists of eight math games appropriate for children in the first grade. The average number of days attended by the students in the comparison group was 45.8 of a possible 50 days. |
| **Primary outcomes and measurement** | For both pre- and posttests, the authors administered four subtests of the Comprehensive Test of Phonological Processing (CTOPP): Blending Words, Blending Non-Words, Elision, and Sound Matching subtests and the Woodcock Reading Mastery Test (WRMT): Word Identification and Word Attack subtests. The CTOPP Memory for Digits subtest and the Spelling subtest of the Wide Range Achievement test were also used in the study, but they have not been included in this review because they are outside the scope of the Beginning Reading review. For a more detailed description of these outcome measures, see Appendices A2.1 and A2.2. |
| **Staff/teacher training** | The experimenter trained the computer lab technicians in each of the three schools on how to use the software. Detailed instructions, attendance sheets, and appropriate rotations of the *Earobics®* games were discussed. No other information on teacher training is provided. |

1. The pretest also confirmed that students' performance was low average.

## Appendix A2.1 Outcome measures for the alphabetics domain

| Outcome measure | Description |
|---|---|
| ***Phonological Awareness*** | |
| **Comprehensive Test of Phonological Processing (CTOPP): Elision subtest** | This subtest measures the ability of a student to manipulate the components of a word. The student is prompted to say a compound word (for example, "cowboy") and then to say the word without the first part "cow," or to say a word without a specific sound (such as "f" in the example of "farm"—"arm") (as cited in Valliath, 2002). |
| **CTOPP: Blending Words subtest** | This subtest measures the ability of a student to combine separately spoken sounds and put them together to form a real word (as cited in Valliath, 2002). |
| **CTOPP: Sound Matching subtest** | This subtest measures the ability of a student to choose the word that contains a target sound. Words are presented orally and the student is shown a card containing pictures of the four words. The student must indicate which word contains the sound. The target sound is tested in both the initial and final positions in the word (as cited in Valliath, 2002). |
| **CTOPP: Blending Non-Words subtest** | This subtest measures the ability of a student to combine sounds that are presented orally and put the separate sounds together to form a nonsense word (as cited in Valliath, 2002). |
| **Dynamic Indicators of Basic Early Literacy Skills (DIBELS): Initial Sounds Fluency subtest** | After presenting the student with four line drawings of nouns, randomly ordered on a sheet of paper, and naming each drawing, the examiner asks the student to identify the letter sounds that each picture begins with (as cited in Rehmann, 2005). |
| **DIBELS: Phoneme Segmentation Fluency subtest** | For this subtest, the examiner reads 24 three- to four-phoneme words and asks the student to say all the sounds they hear in each word. The score is the number of phonemes given in one minute (as cited in Rehmann, 2005). |
| **ORAL-J: Early Literacy Achievement Blending into Words subtest** | This task requires students to combine or blend the separate sounds of a word to say the word. For example, the student is given sounds such as /k/ /a/ /t/ and has to say "cat" (as cited in author communication).[1] |
| **ORAL-J: Early Literacy Achievement Segmenting into Sounds subtest** | This task requires students to segment words into sounds. The student is given a word and has to give individual sounds (as cited in author's communication).[1] |
| **ORAL-J: Early Literacy Achievement: Rhyming Words subtest** | This task requires students to generate words that rhyme. The student is given a word and has to supply a word that rhymes (as cited in author's communication).[1] |

*(continued)*

## Appendix A2.1 Outcome measures for the alphabetics domain *(continued)*

| Outcome measure | Description |
|---|---|
| ***Letter Knowledge*** | |
| **DIBELS: Letter Naming Fluency subtest** | The examiner displays two sets of randomly arrayed letters (uppercase and lowercase) with 10 letters per line, and the student must name as many letters as possible within one minute (as cited in Rehmann, 2005). |
| **ORAL-J: Early Literacy Achievement: Letter Naming subtest** | Students get a card with 100 letters and are to name each one. The subtest score is determined by how many letters they name in one minute (as cited in author's communication).[1] |
| ***Phonics*** | |
| **ORAL-J: Early Literacy Achievement: Sound of Letters subtest** | Students name the sound of letters on a card with 59 letters. The subtest score is determined by how many sounds they name in one minute (as cited in author's communication).[1] |
| **Woodcock Reading Mastery Test (WRMT): Word Identification subtest** | This subtest measures basic word-reading skills and requires the student to read aloud isolated words that range in frequency and difficulty (as cited in Valliath, 2002). |
| **WRMT: Word Attack subtest** | This subtest measures the student's ability to apply phonic and structural analysis skills to pronounce unfamiliar words. Subjects cannot read the pseudowords by sight and must rely on phonologically based processes to decode them (as cited in Valliath, 2002). |
| **DIBELS: Nonsense Words Fluency subtest** | In this subtest, the examiner presents the student with a sheet of paper containing 80 randomly arrayed one-syllable words. The student must read or say the sounds in each word and receives one point for each correct sound within one minute (as cited in Rehmann, 2005). |

1. The information was received from M. Poblanz, author of Cognitive Concepts (2003).

## Appendix A2.2 Outcome measure for the reading fluency domain

| Outcome measure | Description |
|---|---|
| **Dynamic Indicators of Basic Early Literacy Skills (DIBELS): Oral Reading Fluency** | Oral Reading Fluency is a measure of accuracy and fluency with connected text. Students are presented with a passage calibrated at their grade level and asked to read aloud for one minute. Scoring is based on mispronunciations, omissions, substitutions, and hesitations (as cited in Gale, 2006). |
| **ORAL-J: Early Literacy Achievement: Words per Minute (WPM)** | As students read a story, the teacher times their word recognition for one minute. There are three different stories on each ORAL-J test from which three different WPM scores are derived (as cited in author's communication).[1] |

1. The information was received from M. Poblanz, author of Cognitive Concepts (2003).

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

| Outcome measure | Study sample | Sample size (students) | Authors' findings from the study: Mean outcome (standard deviation)[2] — *Earobics®* group | Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group | WWC calculations: Mean difference[3] (*Earobics®* – comparison) | WWC calculations: Effect size[4] | WWC calculations: Statistical significance[5] (at $\alpha = 0.05$) | WWC calculations: Improvement index[6] |
|---|---|---|---|---|---|---|---|---|
| **Cognitive Concepts, 2003 (randomized controlled trial)[7]** | | | | | | | | |
| ***Construct: Phonological Awareness*** | | | | | | | | |
| ORAL-J: Blending into Words subtest[8] | K–Grade 3 | 74 | 17.31[9] (3.54) | 14.86 (4.10) | 2.45 | 0.64 | Statistically significant | +24 |
| ORAL-J: Segmenting into Sounds[8] | K–Grade 3 | 74 | 45.31[9] (14.31) | 35.80 (15.82) | 9.51 | 0.63 | Statistically significant | +23 |
| ORAL-J: Rhyming words[8] | K-Grade 3 | 74 | 7.16[9] (5.31) | 4.26 (4.36) | 2.90 | 0.59 | Statistically significant | +22 |
| ***Construct: Letter Knowledge*** | | | | | | | | |
| ORAL-J: Letter Naming[8] | K–Grade 3 | 74 | 57.49[9] (18.78) | 57.26 (20.63) | 0.23 | 0.01 | ns | 0 |
| ***Construct: Phonics*** | | | | | | | | |
| ORAL-J: Sound of Letters[8] | K–Grade 3 | 74 | 27.80[9] (6.89) | 26.17 (7.72) | 1.63 | 0.22 | ns | +9 |
| **Average for alphabetics (Cognitive Concepts, 2003)[10]** | | | | | | **0.42** | **ns** | **+16** |
| **Gale, 2006 (randomized controlled trial)[7,11]** | | | | | | | | |
| **Comparison #1: *Earobics®* vs. control** | | | | | | | | |
| ***Construct: Phonological Awareness*** | | | | | | | | |
| DIBELS: Initial Sounds Fluency | Kindergarten | 26 | 13.72 (4.61) | 5.21 (3.00) | 8.51 | 2.12 | Statistically significant | +49 |
| DIBELS: Phoneme Segmentation Fluency | Kindergarten | 26 | 1.31[9] (0.75) | 0.00[12] (0.00) | 1.31 | 2.39 | Statistically significant | +49 |
| DIBELS: Phoneme Segmentation Fluency | Grade 1 | 25 | 47.75 (8.08) | 31.02 (10.57) | 16.73 | 1.73 | Statistically significant | +46 |

*(continued)*

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

| Outcome measure | Study sample | Sample size (students) | Authors' findings from the study: Mean outcome (standard deviation)[2] — *Earobics®* group | Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group | WWC calculations: Mean difference[3] (*Earobics®* – comparison) | WWC calculations: Effect size[4] | WWC calculations: Statistical significance[5] (at $\alpha = 0.05$) | WWC calculations: Improvement index[6] |
|---|---|---|---|---|---|---|---|---|
| ***Construct: Letter Knowledge*** | | | | | | | | |
| DIBELS: Letter Naming Fluency | Kindergarten | 26 | 19.69[9] (11.74) | 13.08[12] (10.00) | 6.61 | 0.59 | ns | +23 |
| DIBELS: Letter Naming Fluency | Grade 1 | 25 | 50.26 (13.83) | 38.02 (8.97) | 12.24 | 1.01 | ns | +35 |
| ***Construct: Phonics*** | | | | | | | | |
| DIBELS: Nonsense Words Fluency | Grade 1 | 25 | 47.72 (19.65) | 26.11 (11.44) | 21.61 | 1.29 | Statistically significant | +41 |
| **Average for alphabetics, Comparison #1 (Gale, 2006)[10]** | | | | | | **1.52** | **Statistically significant** | **+44** |
| **Comparison #2: *Earobics®* vs. Lexia** | | | | | | | | |
| ***Construct: Phonological Awareness*** | | | | | | | | |
| DIBELS: Initial Sounds Fluency | Kindergarten | 26 | 13.72 (4.61) | 10.07 (5.01) | 3.65 | 0.73 | ns | +28 |
| DIBELS: Phoneme Segmentation Fluency | Kindergarten | 26 | 1.31[9] (0.75) | 1.31[12] (0.63) | 0.00 | 0.00 | ns | 0 |
| DIBELS: Phoneme Segmentation Fluency | Grade 1 | 25 | 47.75 (8.08) | 37.66 (13.71) | 10.09 | 0.88 | ns | +32 |
| ***Construct: Letter Knowledge*** | | | | | | | | |
| DIBELS: Letter Naming Fluency | Kindergarten | 26 | 21.08[9] (11.74) | 17.31[12] (12.91) | 2.77 | 0.22 | ns | +9 |
| DIBELS: Letter Naming Fluency | Grade 1 | 25 | 50.26 (13.83) | 48.11 (14.33) | 2.15 | 0.15 | ns | +6 |
| ***Construct: Phonics*** | | | | | | | | |
| DIBELS: Nonsense Words Fluency | Grade 1 | 25 | 47.72 (19.65) | 40.87 (15.12) | 6.85 | 0.38 | ns | +15 |
| **Average for alphabetics, Comparison #2 (Gale, 2006)[10]** | | | | | | **0.40** | **ns** | **+16** |
| **Average for alphabetics, Entire study (Gale, 2006)[10]** | | | | | | **0.96** | **ns** | **+34** |

*(continued)*

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

| Outcome measure | Study sample | Sample size (students) | Authors' findings from the study: Mean outcome (standard deviation)[2] — *Earobics®* group | Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group | WWC calculations: Mean difference[3] (*Earobics®* – comparison) | WWC calculations: Effect size[4] | WWC calculations: Statistical significance[5] (at $\alpha = 0.05$) | WWC calculations: Improvement index[6] |
|---|---|---|---|---|---|---|---|---|
| **Rehmann, 2005 (randomized controlled trial)[7, 13]** | | | | | | | | |
| ***Construct: Phonological Awareness*** | | | | | | | | |
| DIBELS: Initial Sounds Fluency | Kindergarten | 30 | 17.3[9] (11.7) | 15.1 (13.6) | 2.20 | 0.17 | ns | +7 |
| DIBELS: Phoneme Segmentation Fluency | Grade 1 | 35 | 37.2[9] (19.5) | 34.6 (18.2) | 2.60 | 0.14 | ns | +5 |
| ***Construct: Letter Knowledge*** | | | | | | | | |
| DIBELS: Letter Naming Fluency | Kindergarten | 31 | 26.7[9] (17.3) | 26.6 (14.3) | 0.10 | 0.01 | ns | 0 |
| ***Construct: Phonics*** | | | | | | | | |
| DIBELS: Nonsense Words Fluency | Grade 1 | 35 | 38.6[9] (20.3) | 48.1 (30.8) | –9.50 | –0.35 | ns | –14 |
| **Average for alphabetics (Rehmann, 2005)[10]** | | | | | | **–0.01** | **ns** | **0** |
| **Valliath, 2002 (quasi-experimental design)[7]** | | | | | | | | |
| ***Construct: Phonological Awareness*** | | | | | | | | |
| CTOPP: Elision | Grade 1 | 30 | 104.00 (11.98) | 97.67 (7.04) | 6.33 | 0.63 | ns | +23 |
| CTOPP: Blending Words | Grade 1 | 30 | 105.66 (4.88) | 103.33 (9.76) | 2.33 | 0.29 | ns | +12 |
| CTOPP: Sound Matching | Grade 1 | 30 | 103.63 (4.58) | 95.00 (9.45) | 8.63 | 1.13 | Statistically significant | +37 |
| CTOPP: Blending Non-Words | Grade 1 | 30 | 111.00 (9.02) | 105.33 (10.26) | 5.67 | 0.57 | ns | +22 |

*(continued)*

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

| Outcome measure | Study sample | Sample size (students) | Authors' findings from the study: Mean outcome (standard deviation)[2] — *Earobics®* group | Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group | WWC calculations: Mean difference[3] (*Earobics®* – comparison) | WWC calculations: Effect size[4] | WWC calculations: Statistical significance[5] (at $\alpha = 0.05$) | WWC calculations: Improvement index[6] |
|---|---|---|---|---|---|---|---|---|
| ***Construct: Phonics*** | | | | | | | | |
| WRMT: Word Identification | Grade 1 | 30 | 104.07 (5.16) | 100.87 (3.76) | 3.2 | 0.69 | ns | +25 |
| WRMT: Word Attack | Grade 1 | 30 | 103.33 (6.18) | 101.33 (6.90) | 2.0 | 0.30 | ns | +12 |
| **Average for alphabetics (Valliath, 2002)[10]** | | | | | | **0.60** | **ns** | **+23** |
| **Domain average for alphabetics across all studies[9]** | | | | | | **0.49** | **na** | **+25** |

ns = not statistically significant
na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the alphabetics domain.
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. 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 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 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. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Cognitive Concepts (2003), Rehmann (2005), and Valliath (2002), no corrections for clustering or multiple comparisons were needed. In the case of Gale (2006), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.
8. Means and standard deviations were received through communication with the author.
9. The *Earobics®* group mean equals the comparison group mean plus the mean difference. The study author did not provide adjusted means for this outcome, so the WWC calculated the mean difference in outcomes, taking into account the pretest difference between the study groups. For further details, please see Technical Details of WWC-Conducted Computations.
10. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.
11. Unless otherwise noted, means are posttest means, ANCOVA-adjusted for pretest differences, as reported in Gale (2006).
12. Unadjusted posttest mean as reported in Gale (2006).
13. Because of severe differential attrition, the WWC requires that post attrition baseline equivalence of the groups be established. The study author used an ANCOVA control for pretest measures to statistically equate the groups for analysis. Therefore, the WWC conducted computations using the means and standard deviations of the analysis sample that remained after attrition occurred, as reported in the study. The author did not provide ANCOVA-adjusted means, so for this study, we calculate the mean difference by adjusting for pretest scores. For further details, please see Technical Details of WWC-Conducted Computations.

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

| Outcome measure | Study sample | Sample size (students) | Authors' findings from the study[2]: Mean outcome (standard deviation)[3]: *Earobics*® group | Authors' findings from the study[2]: Mean outcome (standard deviation)[3]: Comparison group | WWC calculations: Mean difference[4] (*Earobics*®–comparison) | WWC calculations: Effect size[5] | WWC calculations: Statistical significance[6] (at $\alpha = 0.05$) | WWC calculations: Improvement index[7] |
|---|---|---|---|---|---|---|---|---|
| **Cognitive Concepts, 2003 (randomized controlled trial)[8]** | | | | | | | | |
| ORAL-J: Words per Minute 1 | K–Grade 3 | 74 | 39.21[9] (22.95) | 35.49 (26.32) | 3.72 | 0.15 | ns | +6 |
| ORAL-J: Words per Minute 2 | K–Grade 3 | 74 | 34.11[8] (25.91) | 31.63 (33.64) | 2.48 | 0.08 | ns | +3 |
| ORAL-J: Words per Minute 3 | K–Grade 3 | 74 | 36.70[8] (27.35) | 33.86 (32.02) | 2.84 | 0.10 | ns | +4 |
| **Average for reading fluency (Cognitive Concepts, 2003)[10]** | | | | | | **0.11** | **ns** | **+4** |
| **Gale, 2006 (randomized controlled trial)[7, 11]** | | | | | | | | |
| **Comparison #1: *Earobics*® vs. control** | | | | | | | | |
| DIBELS: Oral Reading Fluency | Grade 1 | 25 | 27.35 (18.53) | 13.81 (7.83) | 13.54 | 0.91 | ns | +33 |
| **Comparison #2: *Earobics*® vs. Lexia** | | | | | | | | |
| DIBELS: Oral Reading Fluency | Grade 1 | 25 | 27.35 (18.53) | 21.31 (9.65) | 6.04 | 0.39 | ns | +16 |
| **Average for reading fluency (Gale, 2006)[9]** | | | | | | **0.65** | **ns** | **+25** |
| **Domain average for reading fluency[9]** | | | | | | **0.76** | **na** | **+15** |

ns = not statistically significant

na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the reading fluency domain.
2. Means and standard deviations were received through communication with the author.
3. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.

## Appendix A3.2 Summary of study findings included in the rating for the reading fluency domain[1] *(continued)*

8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Cognitive Concepts (2003), no corrections for clustering or multiple comparisons were needed. In the case of Gale (2006), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.
9. The *Earobics*® group mean equals the comparison group mean plus the mean difference. The study author did not provide adjusted means for this outcome, so the WWC calculated the mean difference in outcomes, taking into account the pretest difference between the study groups. For further details, please see Technical Details of WWC-Conducted Computations.
10. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.
11. Means are posttest means, ANCOVA-adjusted for pretest differences, as reported in Gale (2006).

## Appendix A4.1 *Earobics*® rating for the alphabetics domain

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

For the outcome domain of alphabetics, the WWC rated *Earobics*® as having positive effects. The remaining ratings (potentially positive effects, mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered, as *Earobics*® was assigned the highest applicable 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.** Three studies showed statistically significant positive effects, and two had strong designs. |
| **AND** |
| • Criterion 2: No studies showing statistically significant or substantively important *negative* effects. |
| **Met.** No study showed a statistically significant or substantively important negative effect. |

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

## Appendix A4.2 *Earobics*® rating for the reading fluency domain

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

For the outcome domain of reading fluency, the WWC rated *Earobics*® as having potentially positive effects. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered, as *Earobics*® 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.** Two studies included fluency measures, and one (Gale, 2006) showed a 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.** No study showed a statistically significant or substantively important negative effect. 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 study showed a statistically significant positive effect.

**AND**

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

  **Met.** No study showed a statistically significant effect, and no studies showed 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. For a complete description, see the WWC Intervention Rating Scheme.

## Appendix A5 Extent of evidence by domain

| Outcome domain | Number of studies | Sample size | | Extent of evidence[1] |
|---|---|---|---|---|
| | | Schools | Students | |
| Alphabetics | 4 | 6 | 246 | Small |
| Reading fluency | 2 | 2 | 111 | Small |
| Comprehension | 0 | 0 | 0 | na |
| General reading achievement | 0 | 0 | 0 | na |

na = not applicable

1. A rating of "medium to large" requires at least two studies and two schools across studies in one domain, and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is "small."