Educational Research
Institute of America
www.eriaonline.com • email: info@eriaonline.com
An affiliate of Beck Evaluation & Testing Associates, Inc.

A STUDY OF THE INSTRUCTIONAL EFFECTIVENESS OF *Earobics Reading Program*

Report Number 481

February 2014

Contents

Abstract

The importance of effective beginning reading skills and strategies is crucial if students are to become successful readers. If a child is not able to read by the end of third grade, the odds are great that he or she will never catch up. And the effects of falling behind can be devastating.
High poverty students, students with special needs, students of different cultural backgrounds, English language learners, and urban and rural students all deserve an opportunity to excel.

To help struggling readers in the primary grades develop better reading skills and strategies, Houghton Mifflin Harcourt published, ***Earobics***, a supplemental reading program which provides instruction to primary grade students.

In order to evaluate the program, *Houghton Mifflin Harcourt* contracted with the *Educational Research Institute of America* (ERIA) to conduct a one semester study of the effectiveness of ***Earobics***. The study was conducted during the 2013 first semester.

The program was tried out with three groups of students, those in pre-kindergarten and kindergarten, those in grade 1, and those in grades 2 and 3. The study included ***Earobics*** students in two different schools in 2 different states. The ***Earobics*** program had not been previously used in the schools by any classes. Assessments were developed for each of the three groups based on the specific skills and strategies taught in the program. Material reviews substantiated the content validity of the assessments. Reliability analyses showed that the tests had strong internal consistency and were appropriate for making conclusions about students' reading achievement within the program.

The results showed that the students at all 3 grade level groups made statistically significant gains. The results also showed the ***Earobics*** program proved equally effective with both higher and lower pretest scoring students. For all three groups, the low pretest scoring group made similar or greater gains than the higher pretest scoring group

Overview of the Study

Houghton Mifflin Harcourt contracted with the *Educational Research Institute of America* (ERIA) to conduct a one semester study to determine the effectiveness of ***Earobics*** for elementary school students. The study took place over an academic semester, from September 2013 to December 2013.

Research Questions

The following research questions guided the design of the study and the data analyses:

1. Is ***Earobics*** effective in improving the reading skills and strategies of primary grade students?
2. Is ***Earobics*** effective in improving the reading skills and strategies of lower pretest performing students to the same extent as higher pretest performing students?

Design of the Study

The program's efficacy was evaluated using a pretest/posttest design. The classes include Pre-Kindergarten, Kindergarten, Grade 1, Grade 2, and Grade 3. A total of 12 tryout teachers in 2 different schools in 2 different states were included in the study.

Before the program instruction started, students were administered a comprehensive assessment developed by researchers at ERIA. The assessment was designed to measure each of the major components of the ***Earobics*** program. Pretest and post-test reliability data was calculated for each of the assessments and is presented as part of this report. Pretest and post-test administration was under the direction of the classroom teacher. All tests were returned to ERIA for scoring and analyses.

Program Description

The following focus for the program as put forth by the publisher highlights the importance of a research/best practices based program:

Earobics

Earobics® is a proven research-based intervention solution for raising student achievement in Grades Pre-K–3. Built on Common Core State Standards, Earobics © 2014 provides targeted individualized instruction in Phonemic Awareness, Phonics, Comprehension, Vocabulary, Fluency, and Writing through adaptive technology combined with a robust set of engaging classroom resources, assessment, and Professional Development. Powerful management tools enable teachers and administrators to monitor and chart student progress for differentiation. Earobics can be used with any core reading program.

Earobics *is a multisensory reading intervention solution designed to support at-risk readers and foster a safe and achievement-oriented learning environment. Earobics includes interactive software, guided instruction, student resources, teacher's guides, correlations and assessments, customized professional development, and school-to-home connections.*

Timeline and Program Use

The teachers used ***Earobics*** as a supplementary reading program. Teachers reported using the program 15-20 minutes per day to more than 30 minutes per day. Teachers also reported using the program from 1 to 5 days per week. Teachers reported that time of use depended on students' level of reading skills development. Pretests were administered the middle of September, 2013 and posttests were administered the end of December 2013.

Description of the Research Sample

Tables 1 provides the demographic characteristics of the schools included in the study. It is important to note that the school data does not provide a description of the make-up of the classes that participated in the study. However, the data does provide a general description of the schools and, thereby, an estimate of the make-up of the classes included in the study.

Table 1
Demographic Characteristics
Of the *Earobics* Schools Included in the Study

School	State	Location	Grades	Enrollment	% Minority	% Free/Reduced Lunch	% English Language Learner
1	Georgia	Rural	K-5	608	19%	49%	4%
2	Kentucky	Rural	PK-6	339	3%	68%	N/A
Averages				**474**	**11%**	**58%**	**4%**

Description of the Assessment

The pretest and posttest used in the study were developed to assess pre-reading/reading skills and strategies. The make-up of each of the tests is reported in Table 2.

Table 2
Content Specifications for Tests for Each Group

Pre-Kindergarten/Kindergarten Test Composition				
Rhyming	*Letter Recognition*	*Matching Beginning/Ending Sounds*	*Comprehension Words*	*Total*
8 Points	12 Points	26 Points	14 Points	60 Points

Grade 1 Test Composition					
Sounds: Letters	*Vocabulary*	*Comprehension -Words*	*Comprehension-Sentences*	*Comprehension -Stories*	*Total*
18 Points	15 Points	4 Points	4 Points	10 Points	51 Points

Grade 2-3 Test Composition				
Sounds: Letters	*Vocabulary*	*Word Part Clues*	*Comprehension -Stories*	*Total*
15 Points	15 Points	14 Points	15 points	59 Points

Table 3 provides the statistical results for the administration of the pretests and the post-tests. Both the pretests and the post-tests KR 20 reliabilities indicate the tests were reliable for arriving at decisions regarding the achievement of the students to whom the tests were administered.

Table 3
Pretest and Post-Test Test Statistics *Earobics* Classes

Test	Test	Reliability*	SEM**
Pre-K and K	Pretest	.91	3.38
Pre-K and K	Post-test	.98	1.54
Grade 1	Pretest	.83	2.81
Grade 1	Post-test	.92	2.53
Grade 2-3	Pretest	.90	3.05
Grade 2-3	Post-test	.91	2.91

*Reliability computed using the Kuder-Richardson 20 formula.

** SEM is the Standard Error of Measurement.

Data Analyses

Standard scores were developed in order to provide a more normal distribution of scores. The standard scores were a linear transformation of the raw scores. A mean raw score was translated to a mean standard score of 300 and the standard deviation of the raw scores was translated to 50. Standard scores were then used for the statistical analyses.

Data analyses and descriptive statistics were computed for the standard scores from the assessments. The $\leq$.05 level of significance was used as the level at which increases would be considered statistically significant for all of the statistical tests.

The following statistical analyses were conducted to compare students' pretest scores to posttest scores:

- *Paired Comparison t-Tests* were used to determine whether pretest to post-test gains for the ***Earobics*** students in three grade sets:
    - *Pre-Kindergarten and Kindergarten*
    - *Grade 1*
    - *Grades 2 and 3*
- Each of the three groups was divided into two groups based on pretest scores. *Paired Comparison t-Tests* were used with the group that scored higher and the group that scored lower on the pretest to determine if the ***Earobics*** program was equally effective with lower pretest scoring students as with higher pretest performing students.

An effect-size analysis was computed for each of the paired *t*-tests. Cohen's *d* statistic was used to determine the effect size. This statistic provides an indication of the strength of the effect of the treatment regardless of the statistical significance. Cohen's *d* statistic is interpreted as follows:

.2 = small effect
.5 = medium effect
.8 = large effect

Data Results and Analyses

Grade by Grade Comparisons for *Earobics* Students

A Paired Comparison *t*-Test was used to compare the pretest scores and post-test scores of students in the pre-kindergarten and kindergarten classes. In addition, the total group of pre-kindergarten and kindergarten students was split into two equal sized groups based on the students' pretest scores. Paired Comparison t-Tests were then computed for each group. This was done to determine if the students who scored lower on the pretests made as much improvement as those who scored higher on the pretests.

Pre K and K Comparisons

Table 4 shows that the scores of the pre-kindergarten and kindergarten ***Earobics*** students increased statistically significantly from pretest to post-test, and the effect size was large.

Table 4

Pre K and K Paired Comparison *t*-test Results
Pretest/Posttest Comparison of Standard Scores for
***Earobics* Students**

Test	*Number Students*	*Mean Standard Score*	*SD*	*t-test*	*Significance*	*Effect Size*
Pretest	71	278	45.8	12.433	≤.0001	1.47
Post-test	71	322	44.3			

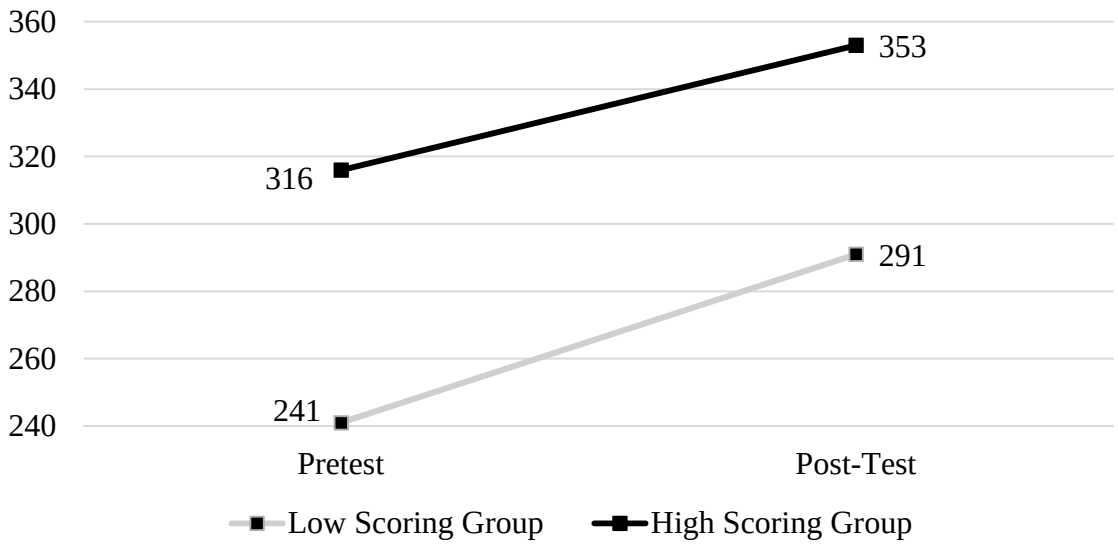
Table 5 shows the results of the Paired Comparison t-tests results for the higher and lower scoring groups. The lower pretest scoring group included a total of 36 students and the higher pretest scoring group included 35 students. The scores of the lower pretesting group averaged 241 and the scores ranged from a low of 191 to a high of 272. The scores of the higher pretesting group averaged 316 and the scores ranged from a low of 272 to a high of 378. Both groups increased statistically significantly and the effect sizes were large for both groups.

Table 5
Pre K and K Paired Comparison *t*-test Results
Pretest/Posttest Comparison of Standards Scores for
High and Low Pretest Scoring *Earobics* Students

Test	*Number Students*	*Mean Standard Score*	*SD*	*t-test*	*Significance*	*Effect Size*
Lower Scoring Students						
Pretest	36	241	19.1	10.020	≤.0001	1.84
Post-test	36	291	34.2			
Higher Scoring Students						
Pretest	35	316	31.2	7.841	≤.0001	1.34
Post-test	35	353	28.9			

Figure 1 shows that the lower group increased at about the same rate as the higher scoring group. The difference between the two groups at pretesting was 75 points and at post-testing it had narrowed slightly to 62 points.

Figure 1
Pre-K and K *Earobics* Students High and Low Pretest Groups



Grade 1 Comparisons

Table 6 shows that the scores of the grade 1 ***Earobics*** students. Their scores increased statistically significantly from pretest to post-test, and the effect size was large.

Table 6
Grade 1 Paired Comparison *t*-test Results
Pretest/Posttest Comparison of Standard Scores for
***Earobics* Students**

Test	*Number Students*	*Mean Standard Score*	*SD*	*t-test*	*Significance*	*Effect Size*
Pretest	30	277	39.0	6.977	≤.0001	1.32
Post-test	30	323	49.6			

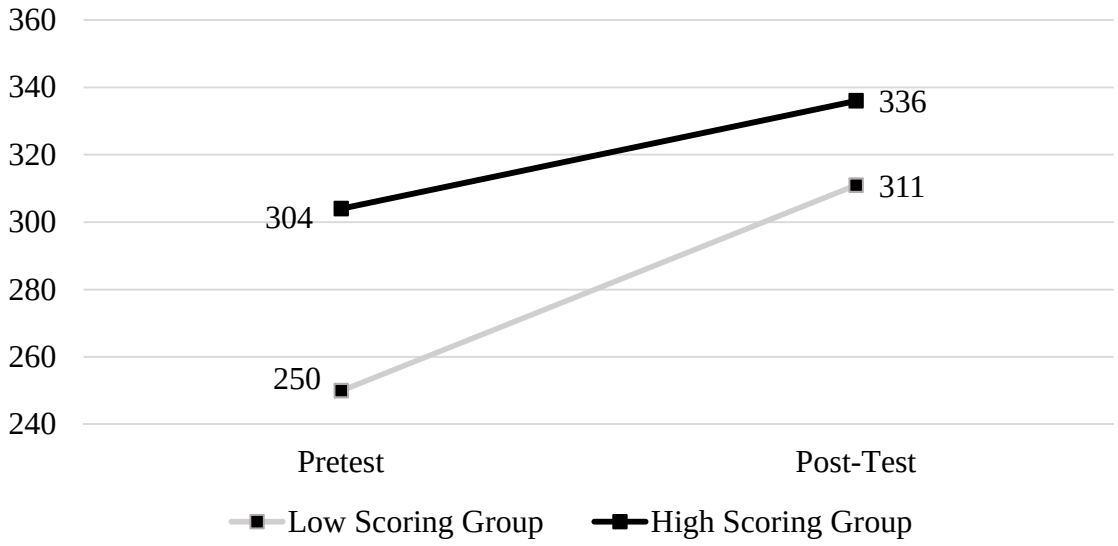
Table 7 shows the results of the Paired Comparison t-tests results for the higher and lower scoring groups. There were a total of 15 students in both groups. The scores of the lower pretesting group averaged 250 and the scores ranged from a low of 155 to a high of 280. The scores of the higher pretesting group averaged 304 and the scores ranged from a low of 286 to a high of 337. Both groups increased statistically significantly and the effect sizes were large for both groups.

Table 7
Grade 1 Paired Comparison *t*-test Results
Pretest/Posttest Comparison of Standards Scores for
High and Low Pretest Scoring *Earobics* Students

Test	*Number Students*	*Mean Standard Score*	*SD*	*t-test*	*Significance*	*Effect Size*
Lower Scoring Students						
Pretest	15	250	36.2	9.419	≤.0001	3.11
Post-test	15	311	52.2			
Higher Scoring Students						
Pretest	15	304	17.0	3.018	≤.009	.97
Post-test	15	336	45.2			

Figure 2 shows that the lower group increased at a higher rate than the higher scoring group. The difference between the two groups at pretesting was 54 points and at post-testing the difference had been cut in half and was only 25 points.

Figure 2
Grade 1 *Earobics* Students High and Low Pretest Groups



Grades 2 and 3 Comparisons

Table 8 shows that the scores of grade 2 and 3 ***Earobics*** students increased statistically significantly from pretest to post-test, and the effect size was medium.

Table 8
Grade 2-3 Paired Comparison *t*-test Results
Pretest/Posttest Comparison of Standard Scores for
***Earobics* Students**

Test	*Number Students*	*Mean Standard Score*	*SD*	*t-test*	*Significance*	*Effect Size*
Pretest	35	290	48.8	3.077	≤.004	.52
Post-test	35	310	49.7			

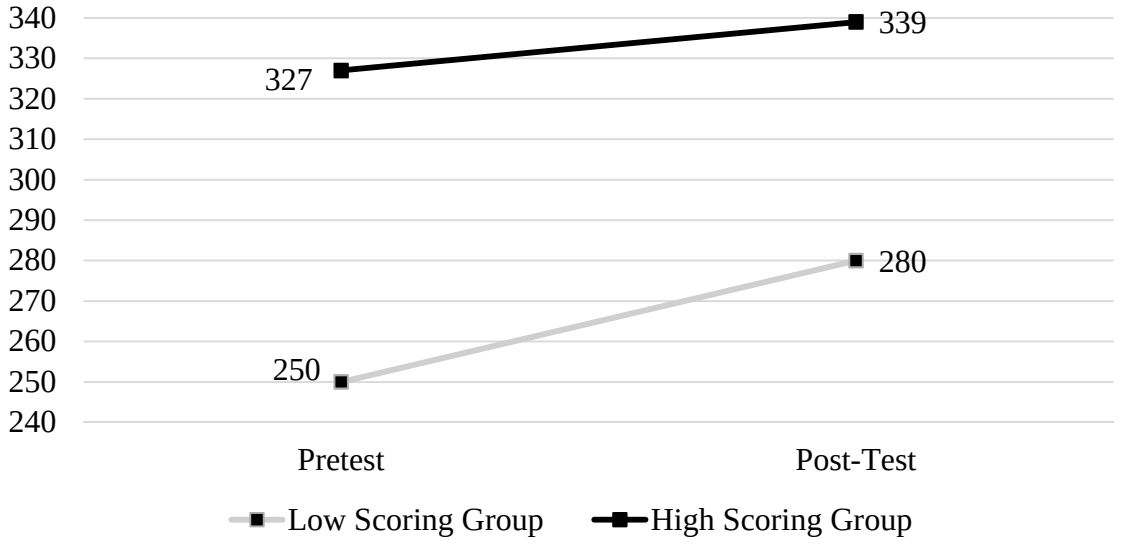
Table 9 shows the results of the Paired Comparison t-tests results for the higher and lower scoring groups. The lower pretest scoring group included a total of 17 students and the higher pretest scoring group included 18 students. The scores of the lower pretesting group averaged 250 and the scores ranged from a low of 167 to a high of 292. The scores of the higher pretesting group averaged 327 and the scores ranged from a low of 292 to a high of 357. Both groups increased statistically significantly and the effect sizes were medium for both the lower scoring group and the higher scoring group.

Table 9
Grade 2-3 Paired Comparison *t*-test Results
Pretest/Posttest Comparison of Standards Scores for
High and Low Pretest Scoring *Earobics* Students

Test	*Number Students*	*Mean Standard Score*	*SD*	*t-test*	*Significance*	*Effect Size*
Lower Scoring Students						
Pretest	17	250	36.3	2.324	≤.03	.58
Post-test	17	280	53.0			
Higher Scoring Students						
Pretest	18	327	23.2	2.423	≤.03	.57
Post-test	18	339	21.8			

Figure 3 shows that the lower group increased at a higher rate than the higher scoring group. The difference between the two groups at pretesting was 77 points and at post-testing it had narrowed to 59 points.

Figure 3
Grade 2-3 Earobics Students High and Low Pretest Groups



Conclusions

This study sought to determine the effectiveness of ***Earobics,*** a supplementary reading program published by Houghton Mifflin Harcourt. The study was carried out with 2 schools in 2 different states. A total of 12 teachers in grades pre-kindergarten, kindergarten, 1, 2, and 3 participated in the study. The tryout teachers were using the program for the first time and received no special instruction in using the program.

Three research questions guided the design and data analyses:

1. Is ***Earobics*** effective in improving the reading skills and strategies of primary grade students?
2. Is ***Earobics*** effective in improving the reading skills and strategies of lower pretest performing students to the same extent as higher pretest performing students?

Question 1: Is ***Earobics*** effective in improving the reading skills and strategies of primary grade students?

Statistical analyses indicated that for all three ***Earobics*** student groups there were statistically significant increases in reading achievement over the academic semester. The effect sizes for these analyses could be categorized as either larger or medium.

Question 2: Is ***Earobics*** effective in improving the reading skills and strategies of lower pretest performing students to the same extent as higher pretest performing students?

The ***Earobics*** students at each grade level were divided into two, approximately equal, groups of those who scored higher on the pretest and those who scored lower. Statistical analyses of students' scores showed that the lower scoring students increased their scores statistically significantly at all grades and the effect sizes were either large or medium. The higher scoring students increased their scores statistically significantly and the effect sizes were either large or medium. These findings indicate that the low performing students made gains as large as, or larger than, the high pretest scoring students.

On the basis of this study, both research questions can be answered positively.

- ***Earobics is effective in improving the reading skills and strategies of primary grade students.***

- ***Earobics is effective in improving the reading skills and strategies of lower pretest performing students to the same or to a greater extent as higher pretest performing students.***