Effectiveness of Houghton Mifflin Harcourt's *Earobics Reach*

A STUDY IN BEAUMONT INDEPENDENT SCHOOL DISTRICT

July 30, 2012

Whitney Hegseth
Valeriy Lazarev
Empirical Education Inc.

Acknowledgements

We are grateful to the people in Beaumont Independent School District for their assistance and cooperation in conducting this research. The research was sponsored by Houghton Mifflin Harcourt which provided Empirical Education Inc. with independence in reporting the results.

About Empirical Education Inc.

Empirical Education Inc. is a Palo Alto, California-based research company that provides rigorous and independent evidence to inform school system decisions. The company brings its expertise in research, data analysis, engineering, and project management to customers that include the US Department of Education, educational publishers, foundations, leading research organizations, and state and local education agencies.

Reference this report: Hegseth, W., Lazarev, V. (2012, July). Effectiveness of Houghton Mifflin Harcourt's Earobics Reach: A Study in Beaumont Independent School District. (Empirical Education Rep. No. Empirical_HMH-6035-FR1-Y1-O.1). Palo Alto, CA: Empirical Education Inc. Retrievable from https://www.empiricaleducation.com/past_research/

Effectiveness of Houghton Mifflin Harcourt's *Earobics Reach*

A Study in Beaumont Independent School District

Table of Contents

Introduction

Houghton Mifflin Harcourt (HMH) contracted with Empirical Education Inc. to conduct a study of the effectiveness of the *Earobics Reach* program in Beaumont Independent School District (Jefferson County, Texas). This report summarizes the findings based on our analysis of data from one PK-5 elementary school in Beaumont during the 2010-2011 school year.

Earobics Reach is a reading intervention program targeting elementary and middle school students who are below grade level in reading. It is designed to provide students with individualized instruction to improve their literacy skills. According to HMH, *Earobics Reach* supplies engaging thematic online content, monitors student performance, and aligns with the daily school curriculum to develop literacy and build skills in phonics, vocabulary, comprehension, and fluency. *Earobics Reach* was piloted by schools in four districts across the United States: Beaumont Independent, Broward County, Rutherford County, and Warren Township, during the 2010-2011 school year. This study reports results from the one school in Beaumont Independent School District that completed both study and program requirements.

Analytic Objectives

The study was designed to address the following question.

Is *Earobics Reach* effective at improving reading achievement for the students in the study school in Beaumont, and if so, what is the effect size?

We used a quasi-experimental comparison group design to compare the reading achievement of students who used *Earobics Reach* during the 2010-2011 school year with reading achievement of similar students who did not use this program, adjusting for the differences in pretest and demographic covariates. The individual student was both the unit of assignment and the unit of analysis in this study. Because the analyzed data came from a single site and for the first year of implementation only, findings from this study should be considered preliminary results that could serve to inform future studies of *Earobics Reach*.

Data

Beaumont Independent School District provided student-level data for 4th and 5th grades in one of its 40 schools. Data were provided for the 2006-2007 through the 2010-2011 school years. The dataset contained student demographic information (e.g. English learner status, ethnicity, socioeconomic and special education status), five consecutive years of student scaled scores on the reading portion of the Texas Assessment of Knowledge and Skills (TAKS), and an indicator of student participation in the *Earobics Reach* program. Students' TAKS reading scores were the outcome measure for the analysis.[1]

Analytic Sample

The dataset included data on 202 students enrolled in grades 4 and 5 in the 2010-2011 school year, the first year of *Earobics* implementation. Fourteen of those students used the *Earobics Reach* program, and

[1] The TAKS is a criterion-referenced test based on Texas Education Agency curriculum. Scores on recent TAKS assessments are generally classified into one of three levels: "Did Not Meet the Standard," "Met the Standard," and "Commended Performance" (Texas Education Agency, n.d.).

the remaining 188 students did not. *Earobics Reach* is targeted generally toward students who perform below grade level. At Beaumont, however, only students with identified disabilities were enrolled in the program. All these students were administered a special version of the reading achievement test (TAKS-M).[2] Thus, the comparison group had to be limited to students with identified disabilities and TAKS-M achievement data, and so consisted of nine students. The data for one *Earobics Reach* student who was administered the regular TAKS test was excluded from the *Earobics Reach* group. The resulting analytic sample consisted of 21 students. We were provided relevant covariates for all 21 students, but pretest data were available for only 16 of the students.

Table 1 presents comparative descriptive statistics for the *Earobics Reach* and comparison groups. Although the *Earobics Reach* and comparison groups are slightly imbalanced, the small size of the analytic sample did not allow for further manipulations to improve the balance.

TABLE 1. COMPOSITION OF THE ANALYTIC SAMPLE

	Earobics Reach group	Comparison group
All students	13	8
Native English speakers	12	8
Students receiving free or reduced priced lunch	11	8
Male students	7	3
Students in grade 4	5	5
Students in grade 5	8	3
Average of TAKS 2010 reading score (pretest; standard deviation in parentheses)	2395 (94)	2204 (215)

Methods

This study used a quasi-experimental comparison group design, whereby inclusion in the comparison group was determined by students' disability status and test type (TAKS-M). The program effect was calculated using linear regression, which adjusted for the differences between the *Earobics Reach* and comparison groups.

Because this study used a small sample, reliable estimates could be obtained only when a small number of covariates were used. At the same time, excluding relevant covariates (i.e. those that are known to influence learning outcomes) would result in a bias of which the size and magnitude is

[2] "The Texas Assessment of Knowledge and Skills–Modified (TAKS–M) is an alternate assessment based on modified academic achievement standards designed for students receiving special education services who meet participation requirements for TAKS–M…According to federal regulations, all students, including those receiving special education services, will be assessed on grade-level curriculum. TAKS–M covers the same grade-level content as the Texas Assessment of Knowledge and Skills (TAKS), but TAKS–M tests have been changed in format (larger font, fewer items per page, etc.) and test design (fewer answer choices, simpler vocabulary and sentence structure, etc.)" (Retrieved from http://www.tea.state.tx.us/student.assessment/techdigest/yr1011.aspx).

impossible to predict. An additional complication with this study's sample was that a large proportion of the students did not have a pretest score. Researchers approached these problems by providing two alternative estimates of the treatment effect. One estimate was obtained using the subsample of students that had pretest data and using the pretest as the only covariate. Another estimate was obtained using the full sample (21 students) and using only student demographic characteristics (gender, grade level, and the free and reduced lunch status) as covariates. Although each of these estimates could be biased, the consistency between the two estimates obtained using different approaches is an indication that the true effect is in the same region.

In addition to achievement and demographic data, researchers also collected implementation data in December 2010 and May 2011. Data collected through teacher background forms and two web-based teacher surveys were used to provide descriptive evidence of the implementation. Researchers deployed the surveys to two *Earobics Reach* teachers and one *Earobics Reach* administrator in the study school in Beaumont.

Results

Statistical Results

Table 2 presents the results of the analysis. The two columns correspond to the two methods described in the preceding section: regression on the smaller subsample with the pretest as the only covariate (method 1), and regression using the full sample (21 students) with student demographic characteristics as covariates (method 2). Both estimates are positive and have low *p* values (below .1 and.05 for methods 1 and 2 respectively) despite the small size of the sample. Together with the consistency between the two estimates demonstrated by a substantial overlap between the confidence intervals, this suggests that the program has a significant positive impact on student reading achievement.

TABLE 2. ESTIMATED EFFECT ON *EAROBICS REACH* STUDENTS

	Method 1	Method 2
Number of students	16	21
Average effect estimate, test score scale	138	154
Standard error of the estimate	72	66
Statistical significance (p value)	0.076	0.032
Confidence interval (95%)	-17 ÷ 293	15 ÷ 293
R squared (proportion of total variance explained by the model)	0.59	0.30
Effect size[a]	.91	1.02

[a]Effect size is calculated using the standard deviation of posttest scores within the analytic sample.

Survey Results

Both teachers who participated in the surveys reported that their special education students used *Earobics Reach* as a supplemental program. Students used the program between thirty minutes per week and thirty minutes per day, depending on the availability of computers in the teachers' classrooms. Teachers reported high levels of student engagement with the program and enjoyment of

learning reading through *Earobics Reach*, stating that students especially liked the program's reward system. Both teachers and the administrator also reported high levels of personal satisfaction and agreement with the program in terms of meeting the needs of struggling readers, especially when compared to other technology programs.

Conclusions

The analysis shows that *Earobics Reach* may have a significant positive impact on reading outcomes for students with disabilities. Teacher and administrator survey responses illuminate usage and satisfaction data that support the finding of a positive impact of *Earobics Reach* on student reading achievement. The small size of the analytic sample does not allow for making definitive conclusions or for performing any analyses of differential impact of the program on student subgroups. Further studies using larger datasets are needed to support the findings of this study.

CAUTIONS FOR INTERPRETING THESE RESULTS

- These results are based on a very small sample and, although statistically significant, may be due to a peculiar composition of the sample. Results therefore cannot be generalized.
- The estimated effect may be biased if the decision to include particular students in the program group (enrollment in *Earobics Reach* program) was based on particular student characteristics which were not reflected in the data.

References

Texas Education Agency. (n.d.). *Student Assessment Division*. Retrieved on July 30, 2012 from http://www.tea.state.tx.us/student.assessment/