

**A STUDY OF THE INSTRUCTIONAL EFFECTIVENESS OF  
On Our Way to English © 2014**

Report Number 486

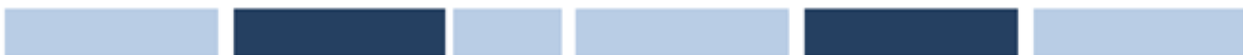
July 2014

**Advisory Board:**

Michael Beck, President  
Beck Evaluation & Testing Associates, Inc.

Jennifer M. Conner, Assistant Professor  
Indiana University

Keith Cruse, Former Managing Director  
Texas Assessment Program



## Contents

---

ABSTRACT.....	2
Overview of the Study .....	3
Research Questions.....	3
Design of the Study.....	4
Timeline and Program Use .....	4
Description of the Research Sample.....	4
Description of the Assessment.....	5
Data Analyses .....	6
Analysis Results.....	7
Grade 2 Analyses .....	7
Higher and Lower Scoring Students.....	7
Grade 4 Analyses .....	9
Higher and Lower Scoring Students.....	9
Conclusions.....	11

## ABSTRACT

---

To help bilingual elementary grade students develop English language skills and knowledge. *Houghton Mifflin Harcourt* has published, *On Our Way to English* © 2014 for students in grades K to 6. *On Our Way to English* includes related visuals, relevant topics and meaningful activities. At the heart of *On Our Way to English* is a commitment to bring rich, culturally relevant language learning to every student.

*Houghton Mifflin Harcourt* contracted with the *Educational Research Institute of America* (ERIA) to conduct a semester long study to test the effectiveness of the program. The study was conducted with students in grades 2 and 4 during the second semester of the 2013/2014 academic year. Sample sizes for both grades were very small because the schools used the program with select groups of students in a pull-out designed instructional program.

Pretests and posttests for both grades 2 and 4 were developed to assess Common Core State Standards-based English language knowledge and skills. Based on these standards a 40 item multiple-choice pretest and post-test assessment was developed for each grade. The assessments focused on the skills, strategies, and knowledge necessary for effective English language understanding and performance.

The results showed that the *On Our Way to English* students made statistically significant gains at both grades 2 and 4 over the course of the semester. The results also showed the *On Our Way to English* program proved effective with both higher and lower pretest scoring students.

The results at both grades were impressive since the program was used for only one semester. In addition, sample sizes were small at both grade levels which often restricts statistically significant gains.

## Overview of the Study

---

This report describes a 2013/2014 one semester study with students in grade 2 and 4 to determine the impact of the *On Our Way to English* © 2014 development program for elementary grade level students. *On Our Way to English* is built upon research and the *Common Core State Standards*. Student learning is based on rich literature and highly visual content support. Both social studies and science units promote academic language skills and vocabulary.

Houghton Mifflin Harcourt school publishers contracted with the *Educational Research Institute of America* (ERIA) to conduct a semester long study to determine the program's effectiveness. The *On Our Way to English* © 2014 was the primary developmental program in the tryout classes.

The development program is described by the publisher on the Houghton Mifflin Harcourt web site as follows:

*On Our Way to English is a comprehensive English language development program that provides everything teachers need for effective instruction. Domain-based instruction includes a focus on academic language and vocabulary development; thematic, content-based instruction; differentiated instruction for language and literacy; and a daily instructional routine in oral language, reading, and writing. Engaging online and digital tools motivate English language learners.*

*The program is designed to engage through visuals, relevant topics and meaningful activities. At the heart of On Our Way to English is a commitment to bring rich, culturally relevant language learning to every English language learner. The instruction is designed to challenge students to reach new heights through rigorous content specifically written to foster success across the four language domains. Built upon the latest research and the Common Core State Standards, On Our Way to English presents language learners with enhanced writing instruction, foundational skills, embedded speaking and listening activities and a myriad of text interaction opportunities that will bring students to the next level of language development.*

## Research Questions

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

1. Is *On Our Way to English* effective in improving the English language knowledge, skills, and understanding of bilingual grade 2 and grade 4 students?

2. Is *On Our Way to English* effective in improving the English language knowledge, skills, and understanding of lower performing as well as higher performing bilingual grade 2 and grade 4 students?

## Design of the Study

The program’s efficacy was evaluated using a pretest/posttest design. The study took place during the second semester of the 2013/2014 academic year in two different schools. Two teachers at grade 2 and 1 teacher at grade 4 participated in the study.

## Timeline and Program Use

The teachers used the *On Our Way to English* text as their primary instructional program. Pretests were administered the end of February 2014 and posttests were administered the middle of June, 2014.

## Description of the Research Sample

Table 1 provides the demographic characteristics of the schools included at each grade level. 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 school and, thereby, an estimate of the make-up of the classes included in the study.

**Table 1**  
**Schools Included in the Study: Demographic Characteristics**

School	State	Location	Grades	Enrollment	% Minority	% Free/Reduced Lunch
<b>Grade 2 Schools Included in the Study: Demographic Characteristics</b>						
1	CT	Urban	PK to 8	1063	97%	100%
2	CT	Urban	PK to 8	900	98%	100%
<i>Average</i>				982	97%	100%
<b>Grade 4 School Included in the Study: Demographic Characteristics</b>						
School	State	Location	Grades	Enrollment	% Minority	% Free/Reduced Lunch
1	CT	Urban	PK to 8	1063	97%	100%

## Description of the Assessment

The pretest and posttest used in the study were developed to assess Common Core State Standards-based language knowledge and skills. Based on these standards a 40 item multiple-choice pretest and post-test assessment was developed for each grade. The assessments focusing on the skills, strategies, and knowledge necessary for effective English language understanding and performance.

Table 2 provides the statistical results for the administration of the pretests and the post-tests administered at each grade. The KR 20 reliability and the Standard Error of Measurement for the assessments at both grades indicates both the pretest score results and the posttest score results were reliable for arriving at decisions regarding the achievement of the students to whom the tests were administered.

**Table 2**  
**Pretest and Post-Test Test Statistics**  
**Grades 2 and 4**

Test	Reliability*	SEM**
<i>Grade 2 Assessments</i>		
Pretest	.85	2.71
Post-test	.86	2.64
<i>Grade 4 Assessments</i>		
Pretest	.82	2.72
Post-test	.83	2.65

\*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 *On Our Way to English* 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:

- A paired comparison *t*-test was used to compare the pretest mean standard scores with the posttest mean standard scores for all students.
- The students were split 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 development program was equally effective with lower performers and higher performers.

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

## Analysis Results

---

### Grade 2 Analyses

Researchers at ERIA conducted a paired comparison *t*-test to determine if the difference from pretest standard scores to posttest standard scores was statistically significant. For this analysis, researchers were able to match the pretest and posttest scores for 38 students. Students who did not take both the pretest and the posttest were not included.

Table 3 shows that the average standard score on the pretest was 285, and the average standard score on the posttest was 315. The increase was statistically significant ( $\leq .0001$ ). The effect size was medium.

**Table 3**  
**Paired Comparison *t*-test Results**  
**Pretest/Posttest Comparison of Standards Scores**

<i>Test</i>	<i>Number Students</i>	<i>Mean Standard Score</i>	<i>SD</i>	<i>t-test</i>	<i>Significance</i>	<i>Effect Size</i>
Pretest	38	285	47.9	5.717	$\leq .0001$	.62
Posttest	38	315	48.2			

### Higher and Lower Scoring Students

An additional analysis was conducted to determine if students who scored lower on the pretest made gains as great as those students who scored higher on the pretest. For this analysis students were ranked in order on the basis of their pretest standard scores. The group of 38 students was divided into two equal sized groups of 19 students. The first group included those students who scored lower on the pretest with a mean of 244 with scores ranging from 186 to 286. The higher scoring group scored an average standard score on the pretest of 326 with scores ranging from 300 to 357.

Pretest-to-posttest comparisons are shown in Table 4 for the lower and higher pretest scoring students. Scores were analyzed using a paired comparison *t*-test to determine if both groups made significant gains.

For both the higher and the lower scoring groups, the average scores increased statistically significantly. The effect size for both the lower scoring and higher pretest scoring groups was large. In line with those results, the data shows that the lower pretest group increased 41 standard score points and the higher pretest scoring group increased 19 standard score points.

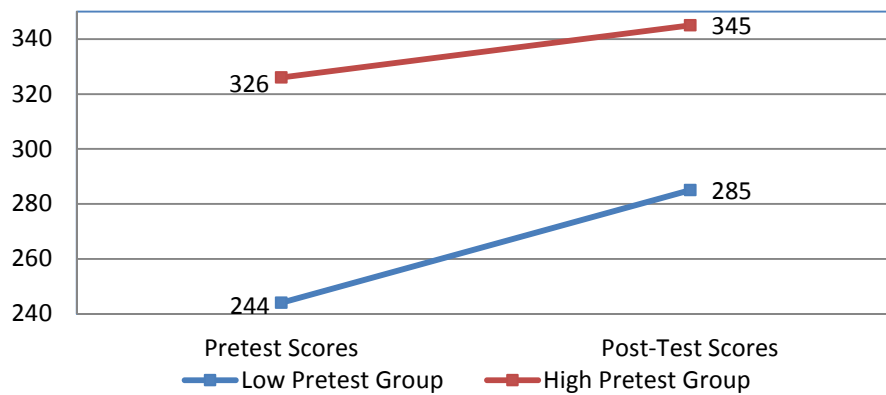


**Table 4**  
**Paired Comparison *t*-test Results for Pretest/Posttest Standard Scores**  
**for the High- and Low-Scoring Pretest Groups**

<i>Test Form</i>	<i>Number Students</i>	<i>Standard Score</i>	<i>SD</i>	<i>t-test</i>	<i>Significance</i>	<i>Effect Size</i>
<b>Lower Scoring Group</b>						
Pretest	19	244	29.2	5.392	≤.0001	1.06
Posttest	19	285	46.1			
<b>Higher Scoring Group</b>						
Pretest	19	326	17.8	2.943	≤.009	.81
Posttest	19	345	27.7			

Figure 1 provides a pretest-to-posttest comparison of the standard scores of lower and higher scoring pretest students. The lower scoring pretest group increased their scores more than the higher scoring pretest group resulting in scores that showed an 82 point difference at the beginning of the academic year and a 60 point difference by the end of the academic year.

**Figure 1**  
**Standard Score Increases for Lower and Higher Pretest Score Students**



## Grade 4 Analyses

Researchers at ERIA conducted a paired comparison *t*-test to determine if the difference from pretest standard scores to posttest standard scores was statistically significant. For this analysis, researchers were able to match the pretest and posttest scores for 20 students. Students who did not take both the pretest and the posttest were not included.

Table 5 shows that the average standard score on the pretest was 282, and the average standard score on the posttest was 318. The increase was statistically significant ( $\leq .0001$ ). The effect size was medium.

**Table 5**  
**Paired Comparison *t*-test Results**  
**Pretest/Posttest Comparison of Standards Scores**

<i>Test</i>	<i>Number Students</i>	<i>Mean Standard Score</i>	<i>SD</i>	<i>t-test</i>	<i>Significance</i>	<i>Effect Size</i>
Pretest	20	282	48.7	9.063	$\leq .0001$	.76
Posttest	20	318	45.8			

### Higher and Lower Scoring Students

An additional analysis was conducted to determine if students who scored lower on the pretest made gains as great as those students who scored higher on the pretest. For this analysis students were ranked in order on the basis of their pretest standard scores. The group of 20 students was divided into two equal sized groups of 10 students. The first group included those students who scored lower on the pretest with a mean of 243 with scores ranging from 215 to 275. The higher scoring group scored an average standard score on the pretest of 321 with scores ranging from 283 to 374.

Pretest-to-posttest comparisons are shown in Table 6 for the lower and higher pretest scoring students. Scores were analyzed using a paired comparison *t*-test to determine if both groups made significant gains.

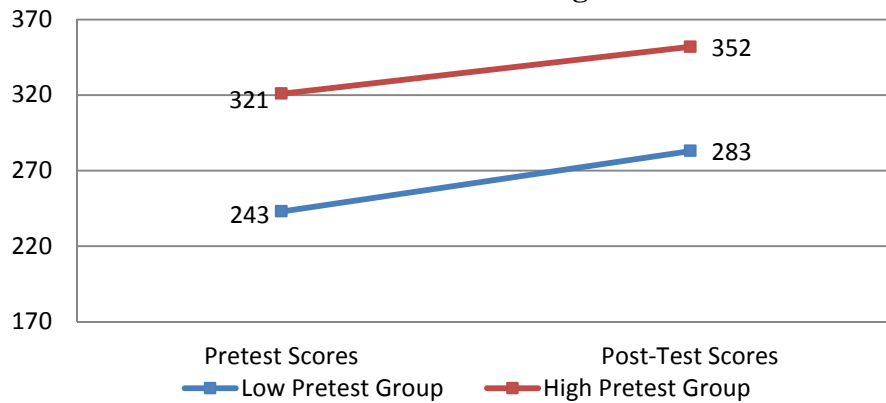
For both the higher and the lower scoring groups, the average scores increased statistically significantly ( $\leq .0001$ ). The effect size for both the lower scoring and higher pretest scoring groups was large. In line with those results, the data shows that the lower pretest group increased 40 standard score points and the higher pretest scoring group increased 31 standard score points.

**Table 6**  
**Paired Comparison *t*-test Results for Pretest/Posttest Standard Scores**  
**for the High- and Low-Scoring Pretest Groups**

<i>Test Form</i>	<i>Number Students</i>	<i>Standard Score</i>	<i>SD</i>	<i>t-test</i>	<i>Significance</i>	<i>Effect Size</i>
<b>Lower Scoring Group</b>						
Pretest	10	243	22.5	6.280	≤.0001	2.07
Posttest	10	283	15.5			
<b>Higher Scoring Group</b>						
Pretest	10	321	32.9	6.997	≤.0001	.86
Posttest	10	352	38.8			

Figure 2 provides a pretest-to-posttest comparison of the standard scores of lower and higher scoring pretest students. The lower scoring pretest group increased their scores more than the higher scoring pretest group resulting in scores that showed a 78 point difference at the beginning of the academic year and a 69 point difference by the end of the academic year.

**Figure 2**  
**Standard Score Increases for Lower and Higher Pretest Score Students**



## Conclusions

---

This study sought to determine the effectiveness of *On Our Way to English* © 2014, a grade K to 5 English language development program published by Houghton Mifflin Harcourt. The study was carried out with classes at grades 2 and 4. The teachers were using the development program for the first time and received no special instruction in using the program.

Two research questions guided the study:

1. Is *On Our Way to English* effective in improving the English language knowledge, skills, and understanding of bilingual grade 2 and grade 4 students?
2. Is *On Our Way to English* effective in improving the English language knowledge, skills, and understanding of lower performing as well as higher performing bilingual grade 2 and grade 4 students?

**Question 1:** Is *Houghton Mifflin Harcourt's On Our Way to English* effective in improving the English language skills and knowledge of elementary school students?

A test designed to assess the English language skills and knowledge was developed to assess students at the beginning and end of a single semester of instruction. Statistical analyses of students' scores at both grades 2 and 4 showed that the students average scores increased statistically significantly. The effect sizes at both grades were medium.

**Question 2:** Is *Houghton Mifflin Harcourt's On Our Way to English* program effective in improving the English language skills and knowledge of lower performing as well as higher performing students at grades 2 and 4?

Statistical analyses of both grade 2 and 4 higher and lower pretest scoring students' both the lower and higher pretest scoring students average student scores increased statistically significant. At both grades for both the lowest pretest scoring students and the higher pretest scoring students the effect size was large.

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

- *The Houghton Mifflin Harcourt's On Our Way to English program is effective in improving the English language skills and knowledge of bilingual grade 2 and grade 4 students.*
- *The Houghton Mifflin Harcourt's On Our Way to English program is effective in improving the English language skills and knowledge of lower performing as well as higher performing bilingual grade 2 and grade 4 students.*