

# *Journeys Common Core*

## *An Efficacy Study*

*Houghton Mifflin Harcourt*

*Report Number 530*

*July 2017*

# Table of Contents

---

Abstract .....	1
Overview of the Study .....	2
<i>Research Questions</i> .....	2
<i>Design of the Study</i> .....	2
<i>Program Description</i> .....	2
<i>Description of the Assessments</i> .....	3
<i>Description of the Study Sample</i> .....	3
Data Analyses and Results .....	5
<i>Grade 2 Results</i> .....	5
<i>Grade 3 Results</i> .....	7
<i>Grade 4 Results</i> .....	8
Conclusions .....	11
<i>Research Question 1</i> .....	11
<i>Research Question 2</i> .....	11

## Abstract

---

The focus of this study was the effectiveness of Journeys Common Core © 2017 a reading program for kindergarten to grade 6 students published by Houghton Mifflin Harcourt. The study included students from 17 different schools in 6 different states. The overall demographics of the study sample are lower than the demographics of students enrolled in public schools in the United States in terms of students eligible for free/reduced lunch programs and below national averages for non-Caucasian students.

The study was conducted with over 1,300 students enrolled in grades 2, 3, and 4. Only those students who took both a pretest and post-test were included in the data analysis. Teachers used the program for their reading instruction five days per week and more than 55 minutes per day. The program was reported as being used by the teachers for the first time. All of the teachers had at least five years of teaching experience and most had 10 to 15 years of teaching experience.

The study was a full year study using the entire Journeys Common Core program with all students as the primary curriculum materials for teaching reading. Pretests and post-tests were developed by reading/language arts curriculum specialists and were based on the Common Core State Standards which were the standards for which the program was developed. In addition to analyzing the gain scores for the total group of students at each grade, analyses were conducted separately for higher and lower scoring reading students. Higher and lower scoring students were identified by the students' pretest scores. Those scoring highest on the pretests were designated as the high scoring reading students and those scoring lowest on the pretests were designated as the lower scoring reading students.

The average gain scores for the total group of students at grades 2, 3, and 4 were statistically significant. The effect sizes of all students at grade 2 were large and at grades 3 and 4 the effect sizes were medium.

In addition, the average gain scores for the low and high scoring groups at each grade level were also statistically significant. The effect sizes for the high and low scoring groups were large at grades 2 and 3. At grade 4 the effect size was large for the low scoring students and medium for high scoring students. *All the effect sizes at every grade exceeded by a large margin the effect sizes needed to determine a substantively important level.*

## Overview of the Study

---

Houghton Mifflin Harcourt, School Publishers contracted with Educational Research Institute of America (ERIA) to conduct a full year study to evaluate the effectiveness of the Journeys Common Core reading program for grades K to 6. The study compared assessments administered to students mid-September 2016 to assessments administered mid-June 2017.

### Research Questions

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

- Does the implementation of *Journeys Common Core Elementary Reading Program* in grades 2, 3, and 4 lead to improved student reading achievement?
- Does the implementation of *Journeys Common Core Elementary Reading Program* in grades 2, 3 and 4 lead to improved student reading achievement as a function of student ability level?

### Design of the Study

The design of the program called for the implementation of the Journeys Common Core program for grades 2, 3, and 4 students during the 2016–2017 academic year. The schools reported that they had not used the program prior to this time.

A total of 70 teachers in six different states participated in the study. The number of teachers at each grade included:

- Grade 2: 33 teachers
- Grade 3: 17 teachers
- Grade 4: 20 teachers

### Program Description

The publisher describes the program on its web page as follows:

*A realistically paced close reading routine and online tools empower students to read rigorous texts. Using authentic text to anchor the core instruction, Journeys weaves the skills of close reading into a practical routine that is designed to have students read and reread for a variety of purposes, giving students more time to dig deeper into the text. The Student eBook provides tools that promote close reading such as responding to questions at point-of-use, highlighting text, and taking notes online. Journeys Close Reader consumable resources feature the high-quality paired text from each Student Book lesson and instruction in reading, re-reading, note-taking, and text annotation—empowering students to read any rigorous text.*

*Journeys “your way” with an array of powerful, easy-to-use digital tools. From the Teacher Dashboard, use the versatile and fully searchable Journeys resources to customize lessons that engage and inspire students. All resources, from lessons to practice materials, are at teachers’ fingertips.*

## Description of the Assessments

The pretest and post-test used in the study were developed by ERIA curriculum experts. Tests were developed to match the content of the Journeys Common Core as well as to emphasize the Common Core State Standards (CCSS).

The tests were developed to respond to the following emphases:

- The grade 2, 3, and 4 assessments included both narrative and expository passages to be read with questions similar to the thinking strategies emphasized by the CCSS. The assessment also included items to assess phonics, vocabulary, and editing/revising of sentences errors in usage and punctuation.

Table 1 provides the test statistics. The table shows that the reliabilities of both the three pretests and the three post-tests are high and provide adequate stability to assess reading achievement.

**Table 1**  
**Pretest and Post-test Statistics for the Journeys Students**  
**Grades 2, 3, and 4**

Test	Mean Score	Standard Deviation	KR 20	SEm*
Grade 2 Pretest	278	47.1	.81	20.53
Grade 2 Post-test	322	42.1	.82	17.86
Grade 3 Pretest	283	53.3	.83	22.0
Grade 3 Post-test	317	39.8	.84	15.9
Grade 4 Pretest	285	54.6	.81	23.8
Grade 4 Post-test	316	39.3	.76	19.3

\*SEm stands for Standard Error of Measurement.

## Description of the Study Sample

Table 2 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.

The percentage of students classified as minority students (non-Caucasian) ranged from 3% to 99% with an average of 28%. By comparison, approximately 50% of the students enrolled in U.S. public schools were classified as non-Caucasian.<sup>1</sup>

The percentage of students enrolled in free/reduced lunch programs ranged from 0% to 99% and averaged 29% across the sample of schools. By comparison, the reported national average for students enrolled in free/reduced lunch programs in public schools was reported as approximately 48%.<sup>2</sup>

<sup>1</sup> The National Center for Educational Statistics (NCES) reported that for the 2011–2012 school year, 48.1% of public school students were enrolled in free/reduced lunch programs. No free/reduced lunch data were available for the 2012–2013 school year. Also, the NCES reported that for the 2012–2013 school year, 49.8% of public school students were classified as minority (non-Caucasian) students.

**Table 2**  
**Demographic Description of the Schools Included in the Study**

	State	Location	Grades	Enrollment	% Non-Caucasian	% FRLP*
1	IL	Town	PK-5	293	96%	99%
2	IL	Suburban	PK-2	483	11%	14%
3	NY	City	PK-5	332	99%	88%
4	PA	Suburban	2-5	740	22%	15%
5	NE	Suburban	PK-6	475	16%	13%
6	NE	Suburban	PK-6	470	23%	30%
7	NE	Suburban	PK-6	381	35%	46%
8	NE	Suburban	PK-6	264	30%	34%
9	NE	Suburban	PK-6	582	12%	7%
10	NE	Suburban	PK-6	291	16%	12%
11	NE	Suburban	PK-6	358	16%	10%
12	NE	Suburban	PK-6	355	14%	25%
13	NE	Suburban	PK-6	523	13%	5%
14	NE	Suburban	PK-6	541	30%	33%
15	NE	Suburban	PK-6	447	14%	15%
16	WY	Rural	K-5	408	31%	44%
17	OH	Rural	K-6	454	3%	0%
AVERAGES				435	28%	29%

## Data Analyses and Results

Standard scores were used for all data analyses. Raw scores were converted to standard scores with a mean of 300 and a standard deviation of 50. Data analyses and descriptive statistics were computed for the students' standard scores.

For all comparisons, paired comparison *t*-tests were used to determine if differences in pretest and post-test scores were significantly different. The  $\leq .05$  level of significance was used as the level at which differences would be considered statistically significant.

In addition, effect size (Cohen's *d*) was computed for each of the comparisons. This statistic provides an indication of the strength of the effect of the treatment regardless of the statistical significance; interpretations of effect sizes in this report include the following guidelines:

.20 to .49 = small

.50 to .79 = medium

.80+ = large

### Grade 2 Results

Table 3 shows that the average scores of the 602 grade 2 students participating in the study increased at a statistical significant level. The effect size was substantively important and is classified as large.

**Table 3**  
**Grade 2 Total Group Paired Comparison *t*-test Results**  
**Pretest/Posttest Standard Score Comparisons**

	<i>Number Students</i>	<i>Mean Standard Score</i>	<i>SD</i>	<i>t-test</i>	<i>Significance</i>	<i>Effect Size</i>
Pretests	602	278	47.1	25.829	$\leq .0001$	1.001
Post-tests	602	322	42.1			

The total group of 602 grade 2 students was divided into two equal sized groups based on their pretest scores. The 301 students scoring lowest on the pretest were lower reading achievement students while the 301-scoring highest on the pretest were higher reading achievement students.

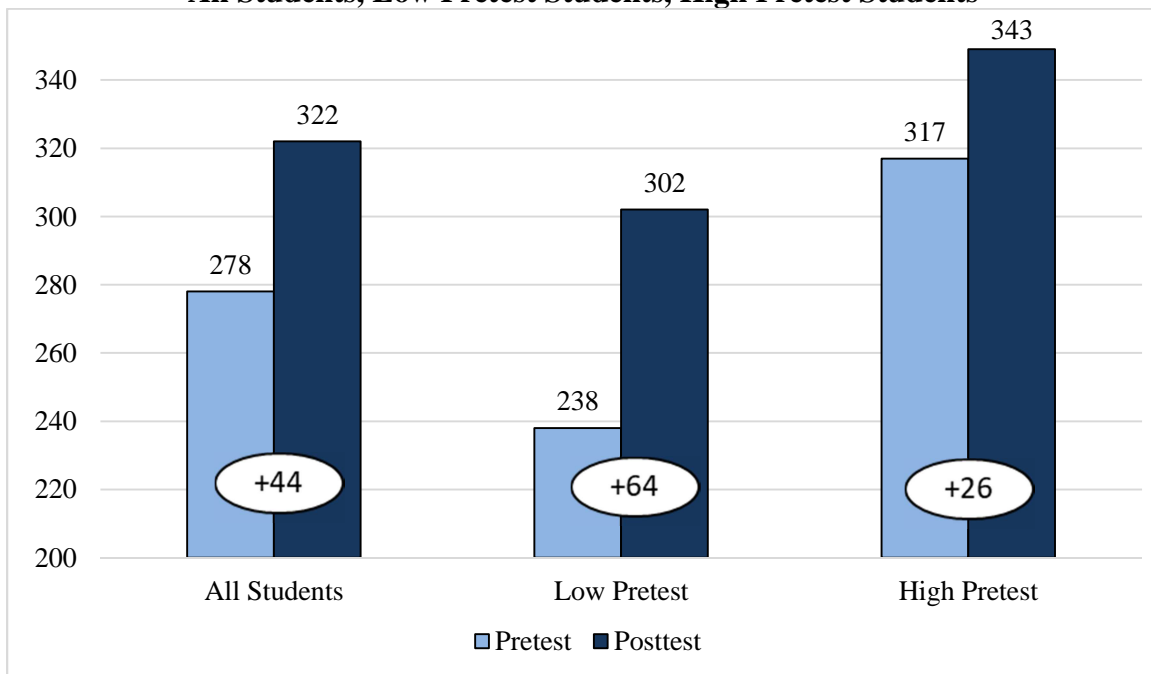
Table 4 shows that both groups made statistically significant gains. The effect sizes for both groups were substantively important and are classified as large.

**Table 4**  
**Grade 2 Paired Comparison *t*-test Results**  
**High- and Low-Scoring Pretest Groups**

<i>Test</i>	<i>Number of Students</i>	<i>Mean Standard Score</i>	<i>SD</i>	<i>t-test</i>	<i>Significance</i>	<i>Effect Size</i>
<b>Lower Scoring Group</b>						
Pretest	301	238	22.9	26.839	≤.0001	1.91
Post-test	301	302	41.6			
<b>Higher Scoring Group</b>						
Pretest	301	317	28.2	12.971	≤.0001	.85
Post-test	301	343	31.9			

Figure 1 provides a graphic representation of the gains achieved by the grade 2 students. In one school year, the grade 2 students increased their average standard scores by 44 standard score points. The low achieving reading students increased their average standard scores by 64 points which was more than twice the increase of the high achieving students whose average standard scores increased by 26 points.

**Figure 1**  
**Grade 2 Pretest/Post-test Gain Comparison**  
**All Students, Low Pretest Students, High Pretest Students**





### Grade 3 Results

Table 5 shows that the average scores of the 360 grade 3 students participating in the study increased their average test scores at a statistical significant level. The effect size was substantively important and is classified as medium.

**Table 5**  
**Grade 3 Total Group Paired Comparison *t*-test Results**  
**Pretest/Post-test Standard Score Comparisons**

	<i>Number Students</i>	<i>Mean Standard Score</i>	<i>SD</i>	<i>t-test</i>	<i>Significance</i>	<i>Effect Size</i>
Pretests	360	283	53.3	16.588	≤.0001	.73
Post-tests	360	317	39.8			

Based on their pretest scores, the total group of 360 grade 3 students was divided into two equal sized groups of 180 students. The students scoring lowest on the pretest were lower reading achievement students while the students scoring highest on the pretest were higher reading achievement students.

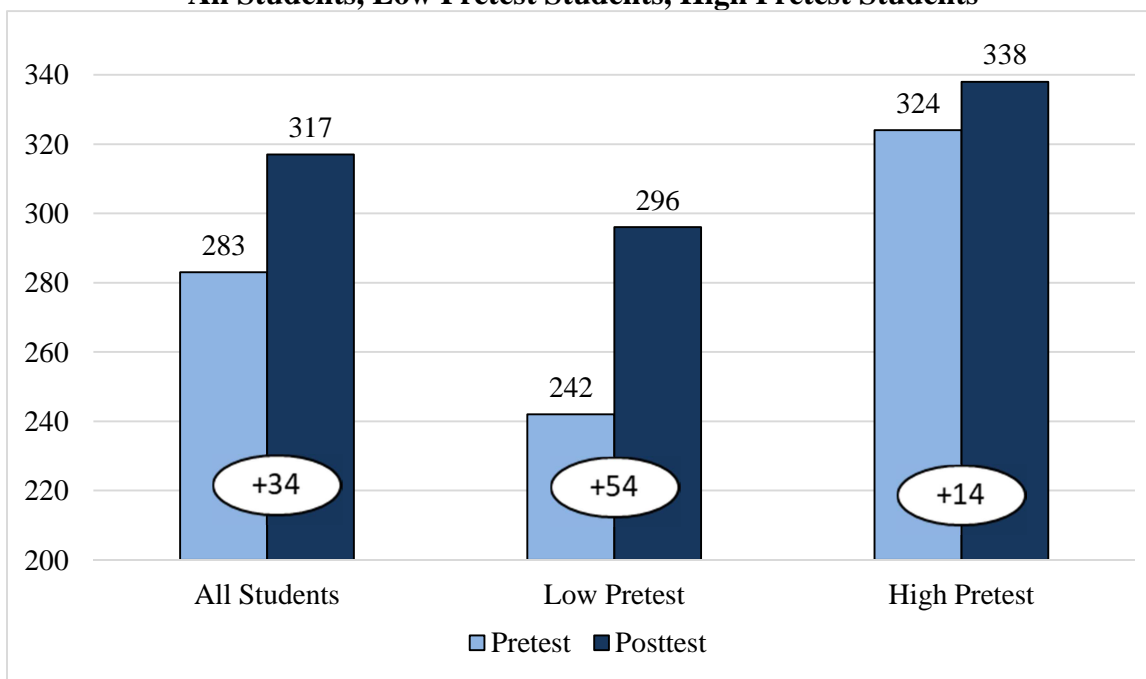
Table 6 shows that both groups made statistically significant gains. The effect sizes for both groups were substantively important and are classified as large.

**Table 6**  
**Grade 3 Paired Comparison *t*-test Results**  
**High- and Low-Scoring Pretest Groups**

<i>Test</i>	<i>Number of Students</i>	<i>Mean Standard Score</i>	<i>SD</i>	<i>t-test</i>	<i>Significance</i>	<i>Effect Size</i>
<b>Lower Scoring Group</b>						
Pretest	180	242	45.3	16.775	≤.0001	1.20
Post-test	180	296	44.9			
<b>Higher Scoring Group</b>						
Pretest	180	324	16.6	9.866	≤.0001	.84
Post-test	180	338	17.1			

Figure 2 provides a graphic representation of the gains achieved by the grade 3 students. In one school year, the grade 3 students increased their average scores by 34 standard score points. The low achieving reading students increased their average scores by 54 standard score points while the high achieving reading students increased their average scores 14 standard score points.

**Figure 2**  
**Grade 3 Pretest Posttest Gain Comparison**  
**All Students, Low Pretest Students, High Pretest Students**



**Grade 4 Results**

Table 7 shows that the average scores of the 362 grade 4 students participating in the study increased their average test scores at a statistical significant level. The effect size was substantively important and is classified as medium.

**Table 7**  
**Grade 4 Total Group Paired Comparison t-test Results**  
**Pretest/Post-test Standard Score Comparisons**

	<i>Number Students</i>	<i>Mean Standard Score</i>	<i>SD</i>	<i>t-test</i>	<i>Significance</i>	<i>Effect Size</i>
Pretests	362	285	54.6	14.589	≤.0001	.65
Post-tests	362	316	39.3			

Based on their pretest scores, the total group of 362 grade 4 students was divided into two equal sized groups of 181 students. The students scoring lowest on the pretest were lower reading achievement students while the students scoring highest on the pretest were higher reading achievement students.

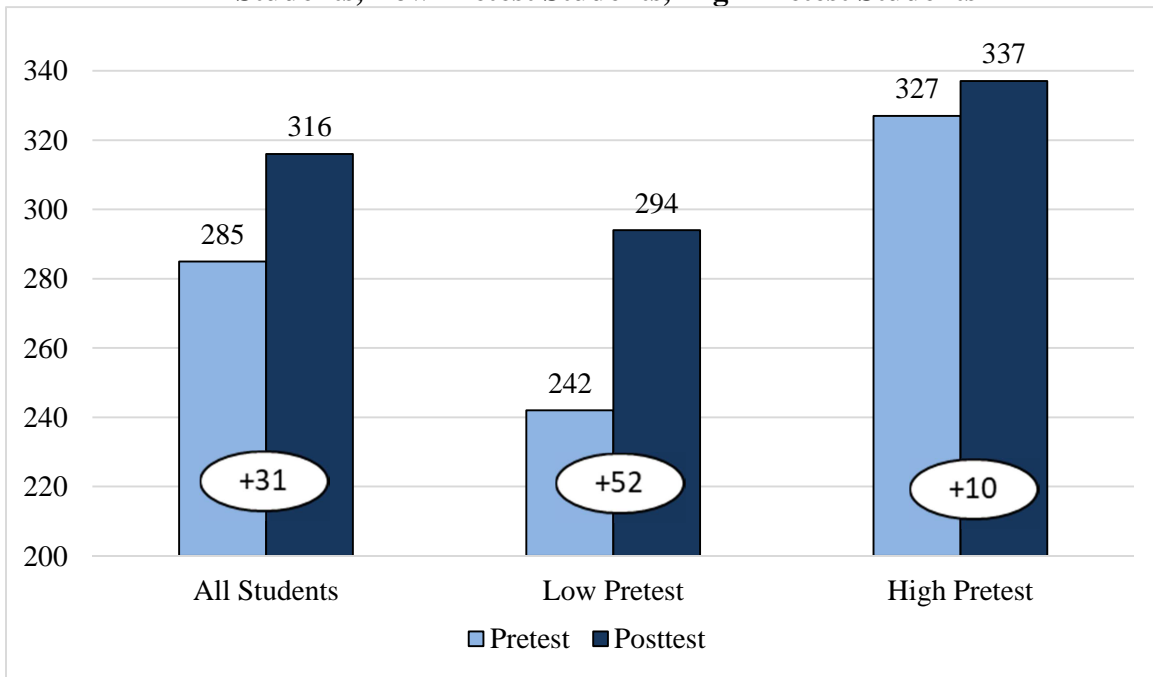
Table 8 shows that both groups made statistically significant gains. The effect sizes for both groups were substantively important and are classified as large for the lower pretest group and medium for the upper pretest scoring group.

**Table 8**  
**Grade 5 Paired Comparison *t*-test Results**  
**High- and Low-Scoring Pretest Groups**

<i>Test</i>	<i>Number of Students</i>	<i>Mean Standard Score</i>	<i>SD</i>	<i>t-test</i>	<i>Significance</i>	<i>Effect Size</i>
<b>Lower Scoring Group</b>						
Pretest	181	242	43.3	17.053	≤.0001	1.25
Post-test	181	294	40.1			
<b>Higher Scoring Group</b>						
Pretest	181	327	21.5	5.505	≤.0001	.56
Post-test	181	337	23.2			

Figure 3 provides a graphic representation of the gains achieved by the grade 4 students. In one school year, the grade 4 students increased their average scores by 31 standard score points. The low achieving reading students increased their average scores by 52 standard score points while the high achieving reading students increased their average scores 10 standard score points.

**Figure 3**  
**Grade 4 Pretest/Post-test Gain Comparison**  
**All Students, Low Pretest Students, High Pretest Students**



## Conclusions

---

This study sought to determine the effectiveness of the Journeys Common Core reading program at grades 2, 3, and 4 by comparing growth on reliable and valid pretests and post-tests. The study took place during the 2016-2017 academic year and was carried out in six states and included 17 different schools and 70 teachers. The student population included a smaller percentage of students eligible for free-reduced price lunch programs than the national average. The percentage of non-Caucasian student was about 25% lower than the national average.

Two research questions guided the study and the conclusions for each are reported below.

### Research Question 1

- Does the implementation of *Journeys Common Core Elementary Reading Program* in grades 2, 3, and 4 lead to improved student reading achievement?

For all three grades included in the study reading achievement growth from pretesting to post-testing was statistically significant. The effect sizes at all grades were above a substantively important level and were large at grade 2 and medium at grades 3 and 4.

### Research Question 2

- Does the implementation of *Journeys Common Core Elementary Reading Program* in grades 2, 3, and 4 lead to improved student reading achievement as a function of student ability level?

For all three grades included in the study reading achievement growth for the high achieving and low achieving students was statistically significant. The effect sizes at grades 2, 3, and 4 for the low pretest scoring students were above a substantively important level and were large at all three grade levels. The effect sizes at grades 2, 3, and 4 for high pretest scoring students were above a substantively important level and were large at grades 2 and 3. The effect size at grade 4 was medium.

Based on this study, both research questions can be answered positively:

The Journeys Common Core program produced statistically significant increases for students at grades 2, 3, and 4. The effect size for the total group at grade 2 were large. At grades 3 and 4 the effect size for the total group of students was medium.

The Journeys Common Core program produced statistically significant growth for both higher ability and lower ability students in grades 2, 3, and 4. The effect sizes for the lower pretest scoring was large at all three grade levels. For the high achieving group, the effect sizes were large at grades 2 and 3 and medium at grade 4.