

*American History*  
*for High School Grades*  
**An Efficacy Study**

*Houghton Mifflin Harcourt*

*Report Number 551*

*June 2018*

## Table of Contents

---

Abstract .....	3
Overview of the Study .....	4
<i>Research Questions</i> .....	4
<i>Design of the Study</i> .....	4
<i>Program Overview</i> .....	4
<i>Description of the Assessments</i> .....	5
<i>Description of the Study Sample</i> .....	6
Data Analyses and Results .....	7
<i>All Student Results</i> .....	7
<i>Results for High and Low Pretest Scoring Students</i> .....	7
Conclusions .....	9

## Abstract

---

The focus of this study was the effectiveness of American History©2018, a social studies program for secondary school students, published by Houghton Mifflin Harcourt. The study included students from three schools in two different states. The overall demographics of the study sample indicate that an average of 43% of the students across the three schools were enrolled in the National School Lunch Program (NSLP) as compared to the national average of 51.8%. The average percentage of non-Caucasian students was 23% which is about half the national average of 48%.

The study was conducted with 188 students most of whom were enrolled in grade 11. Only those students who took both a pretest and posttest were included in the data analysis. The teachers had 5 to ten years of teaching experience and were using the program for the first time in the 2017-2018 school year. The study took place during the second semester. The teachers used the program five days per week and 45 minutes per day.

The study was a second semester study and the teachers agreed that modules 17 through 23 would be used with all students as the primary curriculum materials for teaching social studies. A pretest and post-test were developed by social studies/assessment curriculum specialists and were based on program standards and the content of modules 17 through 23. In addition to analyzing the gain scores for the total group of students at each grade, analyses were conducted separately for higher and lower pretest scoring 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 students and those scoring lowest on the pretests were designated as the lower scoring students.

The average gain score for the total group of students was statistically significant. The effect sizes were also large. In addition, the average gain scores for the low and high pretest scoring groups were statistically significant, and the effect sizes for the high and low scoring groups were large for the low scoring group and medium for the high scoring group.

*All of the effect sizes exceeded by a large margin the effect sizes needed to determine a substantively important gain.*

## Overview of the Study

---

Houghton Mifflin Harcourt school publishers contracted with Educational Research Institute of America (ERIA) to conduct a one semester study to evaluate the effectiveness of the American History social studies program. The study compared program assessments administered to students at the beginning of January 2018 to assessments administered at the end of May 2018.

### Research Questions

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

- Does the implementation of the *American History* program increase students achievement levels to a significant degree?
- Does the implementation of the *American History* program increase students' achievement levels regardless of their achievement level prior to using the program?

### Design of the Study

The design of the program called for the implementation of the American History program for high school students during the 2017–2018 academic year. A total of three teachers from three different schools located in two different states agreed to participate in the study. The teachers reported that the program had not been used in their classes prior to the 2017-2018 academic year. The teachers had five to ten years experience teaching high school history courses. Since the program was new to the teachers using the program, the study was conducted during the second semester.

Most of the students were enrolled in grade 11. The teachers reported using the program 5 days a week with an approximate instructional time of 45 minutes per class period.

### Program Overview

The instructional modules included:

17. The Roaring Twenties
18. The Great Depression
19. The New Deal
20. World War II
21. The Cold War
22. The Postwar Boom
23. An Era of Social Change

The American History program for secondary grade students is described by the publisher as follows:

***Keeping the Story in History***

*Houghton Mifflin Harcourt delivers a powerful and personal narrative of a student’s curiosity and help them connect their learning to their lives and interests.*

***Supporting Inquiry and Active Learning***

*HMH Social Studies offers the tools and support necessary to challenge students to approach history through your inquiry.*

***Providing Choice and Supporting Learning Needs***

*HMH Social Studies American History presents material and activities in a variety of ways to allow students and teachers to choose the path that works best for them. Differentiated instruction and assessments with built-in feedback provide support for all students.*

***Giving the Freedom To Teach Your Way***

*Designed for flexibility, HMH Social Studies American History provides resources in a variety of formats to allow teachers to easily address content in a manner that best fits students’ needs and a teacher’s instructional style.*

***Content Structure***

*In HMH Social Studies American History, the structure of content is shared in digital and print. This enables seamless navigation and content synchronization whether digital, print, or both are used.*

**Description of the Assessments**

The pretest and post-test used in the study were developed by ERIA curriculum assessment experts. Table 1 provides the test statistics. The table shows that the reliabilities of the tests provide adequate stability to assess achievement. Of particular importance is the fact that the test reliabilities are higher for the post-tests than for the pretests. This is almost certainly the result of instruction which would result in less random guessing on the post-tests than on the pretests.

**Table 1  
Pretest and Post-test Statistics  
for the American History Secondary School Students**

<i>Test</i>	<i>Mean Score</i>	<i>Standard Deviation</i>	<i>KR 20</i>	<i>SEm*</i>
Pretest	280	46.2	.58	29.9
Post-test	320	45.3	.71	24.4

*\*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.

An average of 23% percent of students were classified as non-Caucasian. By comparison, approximately 48% of the students enrolled in U.S. public schools were classified as non-Caucasian.

An average of 43% percent of students were enrolled in the National School Lunch programs. By comparison, the reported national average for students enrolled in National School Lunch programs in public schools was reported as approximately 51.8%.<sup>1</sup>

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

State	Location	Grades	Enrollment	Percent		
				Minority	NSLP*	Special Educ.
MO	Suburb: Large	9-12	1,605	36%	49%	6%
UT	Rural: Remote	9-12	324	23%	47%	1%
MO	Town: Distant	9-12	457	9%	33%	N/A
Averages			795	23%	43%	4%

\* NSLP stands for National School Lunch Program

<sup>1</sup> The National Center for Educational Statistics (NCES) reported that for the 2014-2015 school year, 51.8% of public school students were enrolled in free/reduced lunch programs. Also, the NCES reported that for the 2014-2015 school year, 48% of public school students were classified as minority (non-Caucasian) students.

## 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 using the students’ standard scores.

Paired comparison *t*-tests were used to determine if differences in pretest and post-test scores were significantly different. The comparisons were conducted for differences between the *American History January 2018* (pretest) and the *American History May 2018* (post-test). 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. The interpretation of Cohen’s *d* statistic reported by the American Institute for Research (AIR) states that “According to guidelines from the *What Works Clearinghouse*, an effect size of .25 or greater is considered to be ‘substantively important’.” Beyond the level considered to be substantively important, interpretations of effect sizes in this report include the following guidelines:

- .20 to .49 = small
- .50 to .79 = medium
- .80+ = large

### All Student Results

Table 3 shows that the average scores of the 188 students participating in the study increased at a statistical significant level. The effect size was substantively important and was classified as large.

**Table 3**  
**All Students**  
**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	188	280	46.2	11.798	$\leq .0001$	.896
Post-tests	188	320	45.3			

### Results for High and Low Pretest Scoring Students

The total group of 188 students was divided into two equal sized groups based on their pretest scores. The 94 students scoring lowest on the pretest were considered to be lower achieving students while the 94-scoring highest on the pretest scores were considered to be higher achieving students.

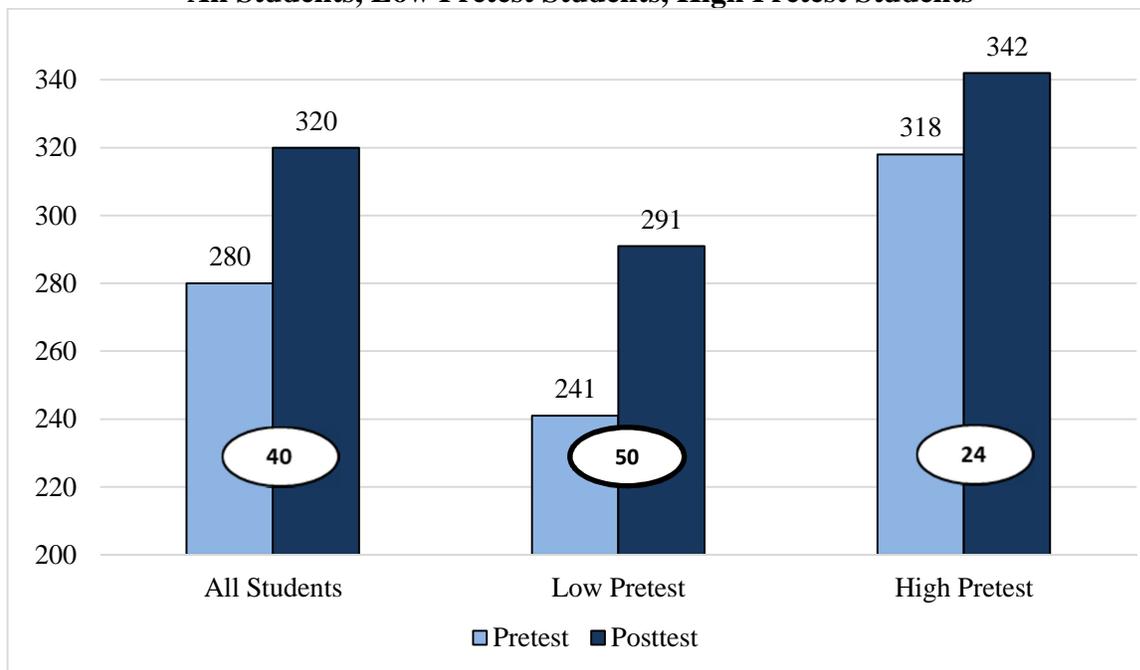
Table 4 shows that both groups made statistically significant gains. The effect sizes for both groups were substantively important and are classified as large for the low pretest scoring group and medium for the high pretest scoring group.

**Table 4**  
**High- and Low-Scoring Pretest Groups**  
**Pretest/Post-test Standard Score Comparisons**

<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	94	241	25.2	12.980	≤.0001	1.783
Post-test	94	291	38.5			
<b>Higher Scoring Group</b>						
Pretest	94	318	25.5	5.051	≤.0001	.697
Post-test	94	342	41.5			

Figure 1 provides a graphic representation of the standard score gains achieved by the students. In one semester, the total group of students increased their average standard scores by 40 standard score points. The low achieving students increased their average standard scores by 50 points which was twice the 24-point increase achieved by the high pretest scoring students.

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



## Conclusions

---

This study sought to determine the effectiveness of the American History program by comparing student's achievement on reliable and valid pretests and post-tests. The study took place during the second semester of the 2017-2018 academic year. The student population included a somewhat smaller percentage of students eligible for free-reduced price lunch programs than the national average. The percentage of non-Caucasian student was only half as large as the national percentage.

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

### *Research Question 1*

- Does the implementation of the *American History* program increase students achievement levels to a significant degree?

For those students included in the study achievement growth from pretesting to post-testing was statistically significant. The effect size was above a substantively important level and was large.

### *Research Question 2*

- Does the implementation of the *American History* program increase student achievement levels regardless of students' achievement levels prior to using the program?

For those students included in the study achievement growth for the high pretest scoring and low pretest scoring students were statistically significant. The effect sizes for both the high and low pretest scoring students were above a substantively important level and were large for the low pretest scoring students and medium for the high pretest scoring students.

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

***The American History program for secondary school students produced statistically significant increases. The effect size for the total group was large.***

***The American History program produced statistically significant growth for both higher pretest scoring students and lower pretest scoring students. The effect size for the higher pretest scoring students was medium and for the lower pretest scoring students the effect size was large.***