

**A STUDY OF THE INSTRUCTIONAL EFFECTIVENESS OF
HOUGHTON MIFFLIN HARCOURT'S
WRITE SOURCE 2012©**

Report Number 396

May 2011

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
Project Background	3
Research Questions	3
Design of the Study	3
Study Participants	4
Timeline and Program Use	4
Description of the Research Sample	5
Description of the Assessment	5
Data Analyses	8
Grade Two Data and Analyses	9
Total Group Analysis	9
Write Source 2012 High and Low Scoring Students.....	10
Grade Four Data and Analyses.....	12
Total Group Analysis	12
Write Source 2012 High and Low Scoring Students.....	13
Grade Seven Data and Analyses	15
Total Group Analysis	15
Write Source 2012 High and Low Scoring Students.....	16
Grade Ten Data and Analyses.....	18
Total Group Analysis	18
Write Source 2012 High and Low Scoring Students.....	19
Conclusions	21

ABSTRACT

*Because of its importance to a student's overall academic success, writing is viewed by educators and the public as particularly important. As a result, educators have become increasingly interested in students' ability to produce effective writing and their knowledge of the skills and strategies of effective writing. To help elementary and secondary students develop better writing skills and strategies, Houghton Mifflin Harcourt has published, **Write Source 2012** 2012©, a program for writing, thinking, and learning.*

In order to evaluate the program's effectiveness, Houghton Mifflin Harcourt contracted with the Educational Research Institute of America (ERIA) to conduct a study to test the effectiveness of the program in several states and school districts.

*The study was conducted in three school districts in three different states. A total of fifteen teachers took part in the study. All of the participating teachers volunteered to participate in the study. The **Write Source 2012** program had not been previously used in the school or district by any classes.*

The study lasted for one semester, the first half of the 2010-2011 school year. A test was designed by ERIA to assess students' writing skills and strategies. A different writing activity was used for the pretest and posttest. The same skills and strategies tests were used as pretest, and after a semester of program, used as a posttest.

*The results showed that the **Write Source 2012** classes made statistically significant gains over the course of the semester. The scores increased significantly for both the Writing Skills and Strategies and for the Writing assessment. The results also showed the **Write Source 2012** program proved equally effective with both higher and lower pretest scoring students.*

Overview of the Study

This report describes a one semester study conducted to determine the impact of *Write Source 2012*®, a writing program for elementary and secondary grade level students. The **Houghton Mifflin Harcourt Write Source 2012** ©program is a resource for teachers who recognize the importance of effective writing instruction.

Project Background

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

Great interest is expressed by parents, legislators, and educators regarding the writing ability of the nation's students. The new **Houghton Mifflin Harcourt Write Source 2012**© program is a resource for teachers who are looking for more effective writing instruction. The activities and strategies in the *Write Source* program are designed to help teachers with the most effective writing instruction. The program is aligned to the Common Core State Standards for English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects. The program designers have based the program on current research and best instructional practice advocated by classroom teachers, administrators, teacher educators, and policymakers alike. The *Write Source* program is planned to provide students with the skills they need to succeed, and prepares them ultimately for college and the workplace. The program helps students develop thinking skills as they choose and develop topics, find information, evaluate sources, think through issues, formulate theses, support arguments, and draw logical conclusions.

Research Questions

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

1. Is **Houghton Mifflin Harcourt's Write Source 2012** ©effective in improving the writing skills and writing performance of elementary and secondary grade level students?
2. Is **Houghton Mifflin Harcourt's Write Source 2012**© effective in improving the writing skills and writing performance of lower performing as well as higher performing elementary and secondary grade level students?

Design of the Study

The program's efficacy was evaluated using a pretest/posttest design. Before program instruction, students were administered a comprehensive pretest designed to cover the content of the chapters included in the study. A similar posttest was used at the end of the study. The difference between the pretests and the posttests was that a different writing topic was used each time. While the writing assignment asked students to write the same type of essay, the specific topic was changed from pretesting to posttesting.

Study Participants

The study was conducted in three school districts in three different states. The program effectiveness data reported here is based on a sample which included the following number of teachers and schools:

Grade 2

2 different schools

5 *Write Source 2012* Classes

Grade 4

2 different schools

5 *Write Source 2012* Classes

Grade 7

2 different schools

3 *Write Source 2012* classes

Grade 10

1 school

2 *Write Source 2012* classes

Timeline and Program Use

All *Write Source 2012* teachers used the program for one semester of the school year of the study. This was the first time the teachers had used the program and most were unfamiliar with the program prior to the tryout.

Description of the Research Sample

Table 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 each 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
Demographic Characteristics
***Write Source 2012* Schools Included in the Study**

Location	Grades Included in the Study	Grades	Students Enrolled	% Students Free/Reduced Lunch Programs	% Minority
Mid-Sized Central City	2	PK to 5	448	35%	54%
Mid-Sized Central City	2	PK to 5	267	96%	90%
Large Central City	4	K to 5	379	95%	100%
Urban Fringe of Large City	4	K to 5	610	15%	40%
Urban Fringe of Large City	7	6 to 8	1450	38%	37%
Urban Fringe of Large City	10	9 to 12	3070	36%	15%

Description of the Assessment

The pretest and posttest used in the study were developed to assess the writing skills and strategies in the *Write Source 2012* instructional materials used in the study. The activities and strategies in the program focus on teaching writing as a process. Students are encouraged to develop a piece of writing over time in five recursive stages that mirror the stages that expert writers go through when working on their own authentic writing. Those five stages include:

- prewriting
- drafting
- revising
- editing/proofreading
- publishing

Based on these teaching goals, a test was developed for each grade level included in the study. The first part of the test, *Skills and Strategies*, was made up of multiple-choice test

items and focused on assessing students' knowledge of types of writing, writing structure and organization, writing mechanics, and editing for correct usage and punctuation. The second part of the test, *Narrative Writing*, asked the student to engage in prewriting and to develop a narrative paragraph.

The *Skills and Strategies* section test was used as both a pretest and posttest. However, the *Narrative Writing* section differed from pretesting to posttesting. A narrative writing activity was included on both the pretest and posttest, but the topic for the writing was changed from pretesting to posttesting.

The outline for each of the tests is described below:

<u>Test</u>	<u>Skills and Strategies</u> <u>Items</u>	<u>Narrative Writing Assessment</u>
Gr 2 Pre	30	...tell about a time that you helped someone.
Gr 2 Post	30	...tell about a time something made you feel proud.
Gr 4 Pre	30	...tell about a time you faced one of your fears.
Gr 4 Post	30	...tell about a time you learned a new skill.
Gr 7 Pre	35	...tell about a time you learned an unexpected lesson.
Gr 7 Post	35	...tell about a time you faced and overcame an obstacle.
Gr 10 Pre	35	...tell about a time you stood up for something you believed in.
Gr 10 Post	35	...tell about a time that something seemed bad but turned out to be positive.

The *Skills and Strategies* tests were scored with one point for each test item. The *Writing Assessments* were scored using the writing rubrics included in the *Write Source 2012* program. Those rubrics included a six-trait model that included four levels for each of the six traits. Only four of the six-traits were included at grade 2. The six writing traits included ideas, organization, voice, word choice, sentence fluency, and conventions. However, word choice and sentence fluency were not included as scoring rubrics for the grade 2 writing assessment.

The Writing assessment were scored using the rubrics which allowed 16 possible points at grade 2 and 24 possible points at grades 4, 7, and 10. All Writing assessments were scored by trained and experienced holistic writing scorers at ERIA. Every tenth paper was scored twice to assure consistency of scoring. Discrepancies in scoring were ameliorated through discussion and the use of a third scorer when differences persisted.

In order to approximate a normal distribution of test scores, the raw scores for the Skills and Strategies test, the Writing test, and the Total test scores were converted to standard scores using a mean of 300 and a standard deviation of 50.

Table 2 provides the test statistics for the multiple-choice test items. The table shows that the reliabilities of all four Skills and Strategies tests are high and provide adequate stability.

Table 2
Kuder Richardson 20 Reliability Estimates, Mean Standard Scores for the
Multiple-Choice Questions on the *Write Source 2012* Posttest Assessment

Test	Mean Standard Score	Standard Deviation	Post Test KR 20
Skills and Strategies Grade 2	319.3	46.4	.77
Skills and Strategies Grade 4	315.9	43.5	.76
Skills and Strategies Grade 7	305.6	55.5	.87
Skills and Strategies Grade 10	308.6	50.0	.76

Data Analyses

Data analyses and descriptive statistics were computed for the standard scores from the **Write Source 2012** assessments at each grade. The $<.05$ 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 at grade 2, grade 4, grade 7, and grade 10:

- 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 at each grade level were split into two groups based on pretest scores. Paired comparison *t*-tests were used with the group that scored highest and the group that scored lowest on the pretest to determine if the program was equally effective with low performers and high performers.

Descriptive statistics were also used to compare pretest and posttest standard test scores at all four grades included in the study.

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

Grade Two Data and Analyses

Total Group Analysis

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:

- Skills and Strategies test
- Writing assessment
- Total score

For this analysis, researchers were able to match the pretest and posttest scores for 87 students. Students who did not take both the pretest and the posttest were not included in the analyses. In addition some students did not record their student numbers legibly or student numbers were missing.

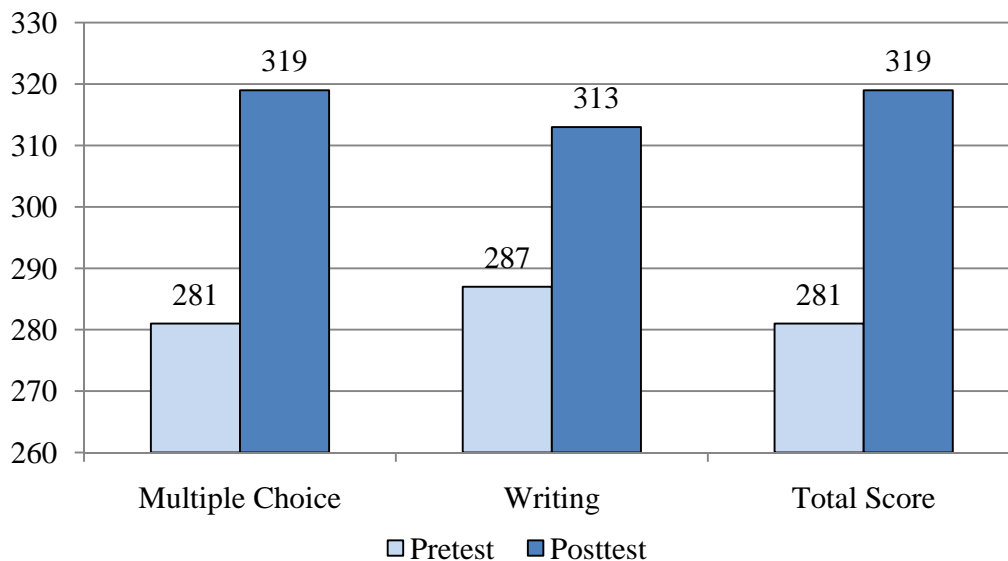
Table 3 shows that all of the gains were statistically significant ($\leq .0001$). The effect size of the gains on the Skills and Strategies questions was large as was the effect size for the Total test score. The effect size for the Writing assessment was medium.

Table 3
Grade 2 Total Paired Comparison t-test Results
Pretest/Posttest Comparison of Standards Scores
for Skills and Strategies, Writing, and Total Scores

Test	Number Students	Mean Standard Score	SD	<i>t</i> -test	Significance	Effect Size
Skills and Strategies (Pre)	87	280.7	44.5	9.161	$\leq .0001$.84
Skills and Strategies (Post)	87	319.3	47.9			
Writing (Pre)	87	286.7	53.2	6.962	$\leq .0001$.55
Writing (Post)	87	313.3	42.9			
Total (Pre)	87	280.9	46.3	11.641	$\leq .0001$.82
Total (Post)	87	319.1	46.4			

Figure 1 provides a graphic look at the changes in standard scores from pretesting to posttesting for the *Write Source 2012* students on the pretests and the posttests.

Figure 1
Grade 2 Standard Scores Pretest and Posttest



***Write Source 2012* High and Low 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 Total *Write Source 2012* standard scores. The group of 87 students was divided into two groups of 43 and 44 students. The first group included those students who scored lower on the pretest. Their average standard score on the pretest was 243 with scores ranging from 189 to 274. The higher scoring group scored an average standard score on the pretest of 318 with scores ranging from 274 to 377.

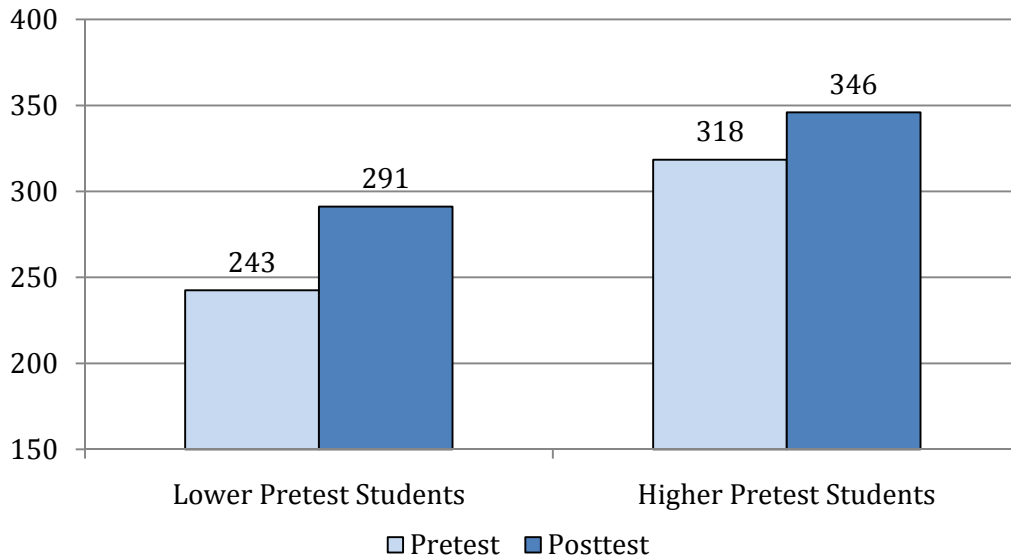
Pretest-to-posttest standard score comparisons are shown in Table 4 for the lower and higher pretest scoring students. For both the higher and the lower scoring group, the total gains were statistically significant ($\leq .0001$). The effect sizes for both groups were large. However, the lower pretest group increased 49 standard score points and the higher pretest scoring group increased 38 standard score points.

Table 4
Grade 2 Paired Comparison *t*-test Results for Pretest/Posttest Standard Scores for the High- and Low-Scoring Pretest Groups

Test	Test Form	Number of Students	Mean Standard Score	SD	t-test	Significance	Effect Size
Lower Scoring Group							
Total	Pretest	43	242.5	24.3	9.736	≤.0001	1.56
Total	Posttest	43	291.3	37.0			
Higher Scoring Group							
Total	Pretest	44	318.5	28.0	7.577	≤.0001	.83
Total	Posttest	44	346.2	38.0			

Figure 2 provides a pretest-to-posttest comparison of the total test standard scores of lower and higher scoring pretest students.

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



Grade Four Data and Analyses

Total Group Analysis

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:

- Skills and Strategies questions
- Writing assessment
- Total score

For this analysis, researchers were able to match the pretest and posttest scores for 69 students. Students who did not take both the pretest and the posttest were not included in the analyses. In addition some students did not record their student numbers legibly or student numbers were missing.

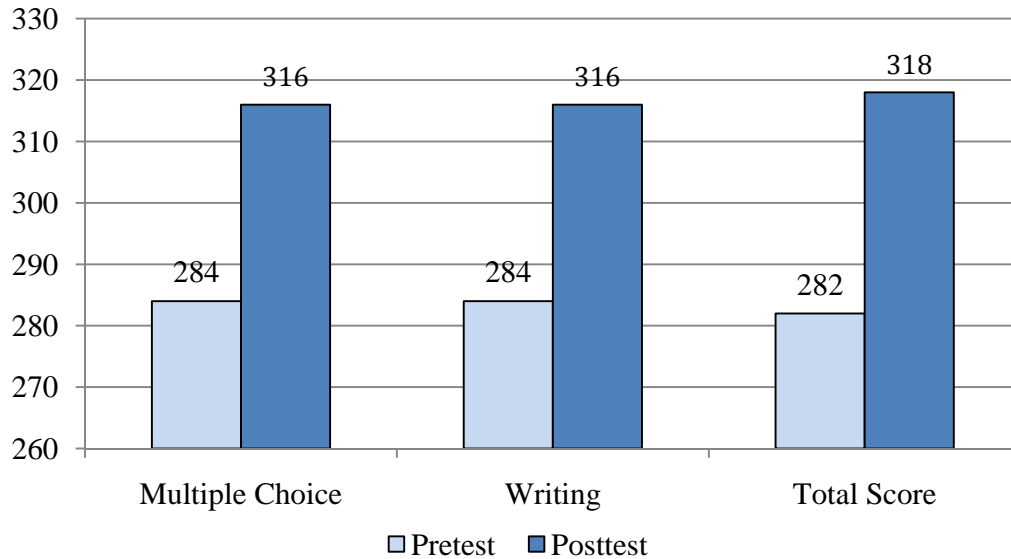
Table 5 shows that all of the gains were statistically significant ($\leq .0001$). The effect size for each of the gains was medium.

Table 5
Grade 4 Total Paired Comparison *t*-test Results
Pretest/Posttest Comparison of Standards Scores
for Skills and Strategies, Writing, and Total Scores

Test	Number Students	Mean Standard Score	SD	<i>t</i> -test	Significance	Effect Size
Skills and Strategies (Pre)	69	284.1	49.2	7.466	$\leq .0001$.67
Skills and Strategies (Post)	69	315.9	45.9			
Writing (Pre)	69	283.5	51.6	7.965	$\leq .0001$.64
Writing (Post)	69	316.5	42.7			
Total (Pre)	69	281.9	49.8	10.223	$\leq .0001$.78
Total (Post)	69	318.1	43.5			

Figure 3 provides a graphic look at the changes in standard scores from pretesting to posttesting for the students on the pretests and the posttests.

Figure 3
Grade 4 Standard Scores Pretest and Posttest



Write Source 2012 High and Low 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 Total ***Write Source 2012*** standard scores. The group of 69 students was divided into two groups of 35 and 34 students. The first group included those students who scored lower on the pretest and their average standard score on the pretest was 243, with scores ranging from 151 to 288. The higher scoring group scored an average standard score on the pretest of 321 with scores ranging from 288 to 381.

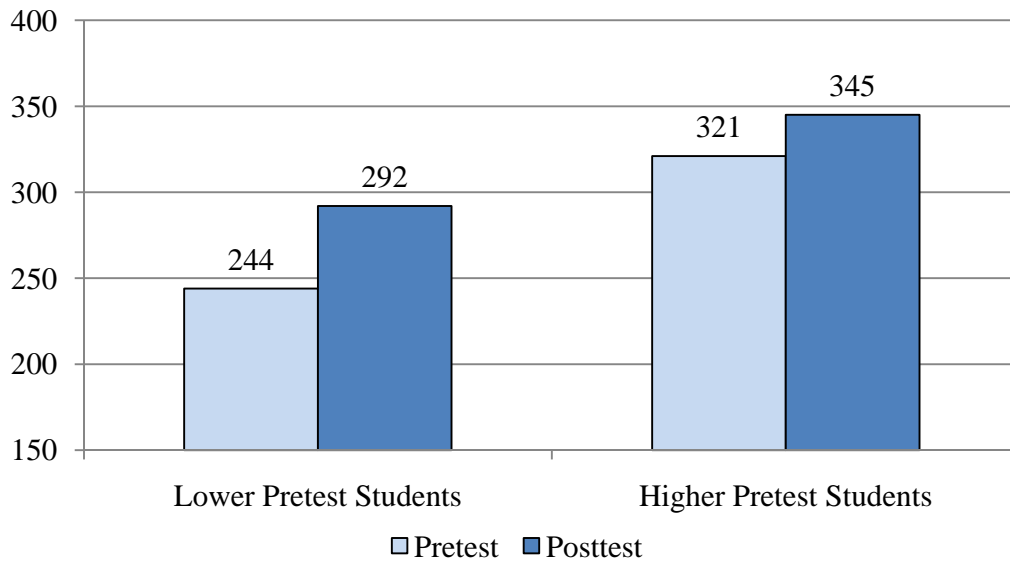
Pretest-to-posttest standard score comparisons are shown in Table 6 for the lower and higher pretest scoring students. For both the higher and the lower scoring group, the total gain was statistically significant ($\leq .0001$). The effect size for both groups was large. However, the lower pretest group increased 48 standard score points and the higher pretest scoring group increased 24 standard score points.

Table 6
Grade 4 Paired Comparison *t*-test Results for Pretest/Posttest Standard Scores for the High- and Low-Scoring Pretest Groups

Test	Test Form	Number of Students	Mean Standard Score	SD	t-test	Significance	Effect Size
Lower Scoring Group							
Total	Pretest	35	243.6	36.0	10.510	≤.0001	1.35
Total	Posttest	35	291.6	35.2			
Higher Scoring Group							
Total	Pretest	34	321.2	25.1	5.183	≤.0001	.82
Total	Posttest	34	345.4	33.3			

Figure 4 provides a pretest-to-posttest comparison of the total test standard scores of lower and higher scoring pretest students.

Figure 4
Grade 4 Standard Score Increases for Lower and Higher Pretest Score Students



Grade Seven Data and Analyses

Total Group Analysis

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:

- Skills and Strategies questions
- Writing assessment
- Total score

For this analysis, researchers were able to match the pretest and posttest scores for 87 students. Students who did not take both the pretest and the posttest were not included in the analyses. In addition some students did not record their student numbers legibly or student numbers were missing.

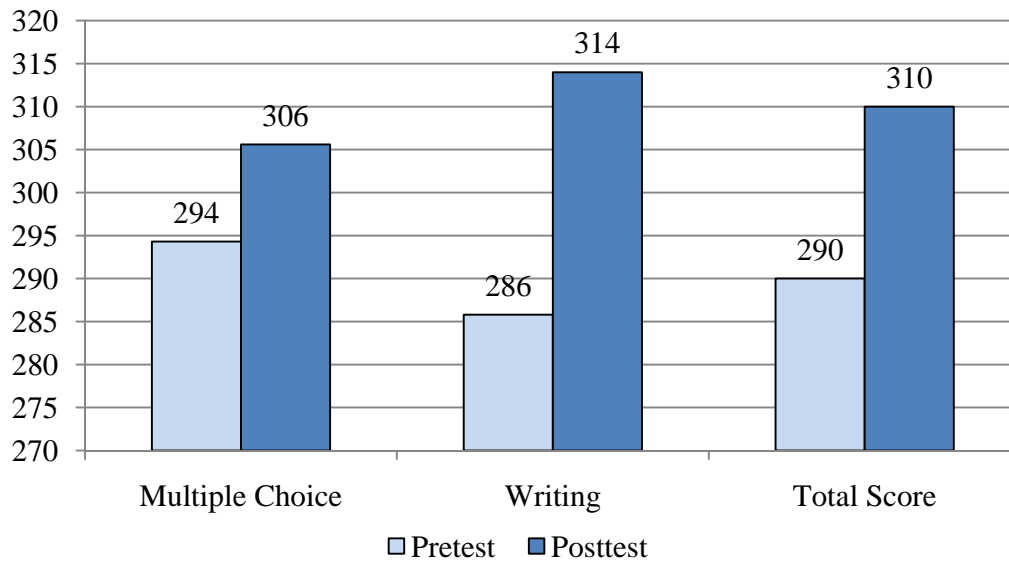
Table 7 shows that all of the gains were statistically significant ($\leq .0001$, $\leq .03$). The effect sizes for the Skills and Strategies questions and the total test score were small while the effect size for the writing test was medium.

Table 7
Grade 7 Total Paired Comparison *t*-test Results
Pretest/Posttest Comparison of Standards Scores
for Skills and Strategies, Writing, and Total Scores

Test	Number Students	Mean Standard Score	SD	<i>t</i> -test	Significance	Effect Size
Skills and Strategies (Pre)	87	294.4	43.4	2.216	$\leq .03$.40
Skills and Strategies (Post)	87	305.6	55.5			
Writing (Pre)	87	285.8	49.4	8.116	$\leq .0001$.59
Writing (Post)	87	314.2	46.7			
Total (Pre)	87	289.6	47.0	5.112	$\leq .0001$.42
Total (Post)	87	310.4	51.0			

Figure 5 provides a graphic look at the changes in standard scores from pretesting to posttesting for the *Write Source 2012* students on the pretests and the posttests.

Figure 5
Grade 7 Standard Scores Pretest and Posttest



***Write Source 2012* High and Low 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 Total *Write Source 2012* standard scores. The group of 87 students was divided into two groups of 43 and 44 students. The first group included those students who scored lower on the Pretest and their average standard score on the Pretest was 252, with scores ranging from 153 to 291. The higher scoring group scored an average standard score on the Pretest of 326 with scores ranging from 291 to 379.

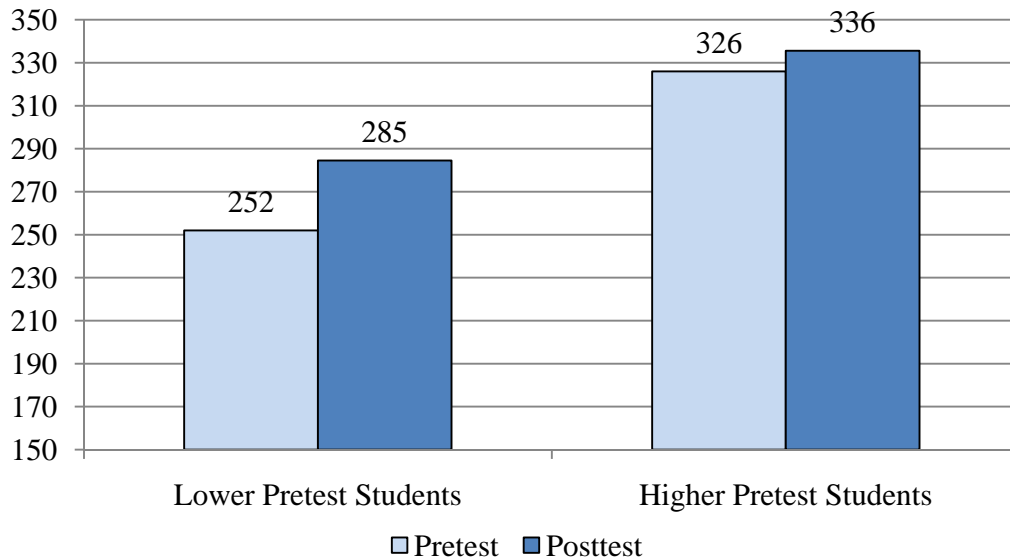
Pretest-to-posttest standard score comparisons are shown in Table 8 for the lower and higher Pretest scoring students. For both the higher and the lower scoring group, the total gain was statistically significant ($\leq .0001$, $\leq .05$). The effect size for the lower scoring group was medium and for the higher scoring group the effect size was small. However, the lower Pretest group increased 32 standard score points while the higher Pretest scoring group increased 10 standard score points.

Table 8
Grade 7 Paired Comparison t-test Results for Pretest/Posttest Standard Scores
for the High- and Low-Scoring Pretest Groups

Test	Test Form	Number of Students	Mean Standard Score	SD	t-test	Significance	Effect Size
Lower Scoring Group							
Total	Pretest	43	252.2	32.0	5.237	≤.0001	.76
Total	Posttest	43	284.5	50.5			
Higher Scoring Group							
Total	Pretest	44	326.2	25.6	1.977	≤.05	.31
Total	Posttest	44	335.6	37.2			

Figure 6 provides a Pretest-to-posttest comparison of the total test standard scores of lower and higher scoring Pretest students.

Figure 6
Grade 7 Standard Score Increases for Lower and Higher Pretest Score Students



Grade Ten Data and Analyses

Total Group Analysis

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:

- Skills and Strategies questions
- Writing assessment
- Total score

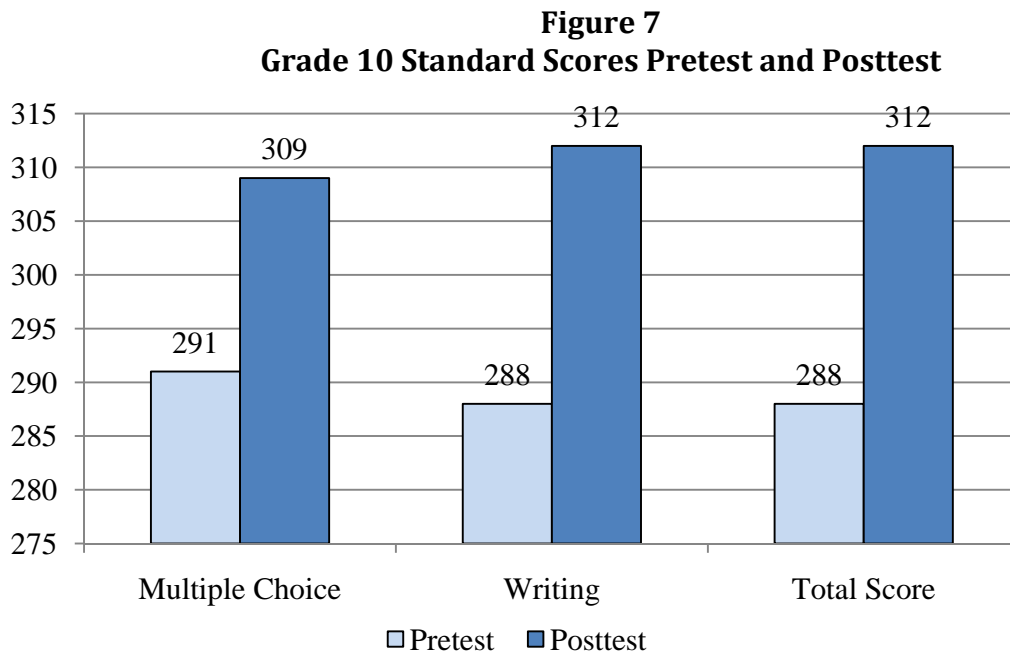
For this analysis, researchers were able to match the pretest and posttest scores for 96 students. Students who did not take both the pretest and the posttest were not included in the analyses. In addition some students did not record their student numbers legibly or student numbers were missing.

Table 9 shows that all of the gains were statistically significant ($\leq .0001$). The effect size for the Skills and Strategies questions was small while the effect size for the writing test and the total test score was medium.

Table 9
Grade 10 Total Paired Comparison t-test Results
Pretest/Posttest Comparison of Standards Scores
for Skills and Strategies, Writing, and Total Scores

Test	Number Students	Mean Standard Score	SD	t-test	Significance	Effect Size
Skills and Strategies (Pre)	87	291.4	48.8	4.196	$\leq .0001$.35
Skills and Strategies (Post)	87	308.6	50.0			
Writing (Pre)	87	288.3	56.4	6.781	$\leq .0001$.51
Writing (Post)	87	311.7	39.6			
Total (Pre)	87	288.2	52.8	6.903	$\leq .0001$.50
Total (Post)	87	311.8	44.3			

Figure 7 provides a graphic look at the changes in standard scores from pretesting to posttesting for the *Write Source 2012* students on the pretests and the posttests.



***Write Source 2012* High and Low 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 Total *Write Source 2012* standard scores. The group of 96 students was divided into two groups of 48. The first group included those students who scored lower on the Pretest and their average standard score on the Pretest was 247, with scores ranging from 140 to 303. The higher scoring group scored an average standard score on the Pretest of 329 with scores ranging from 303 to 406.

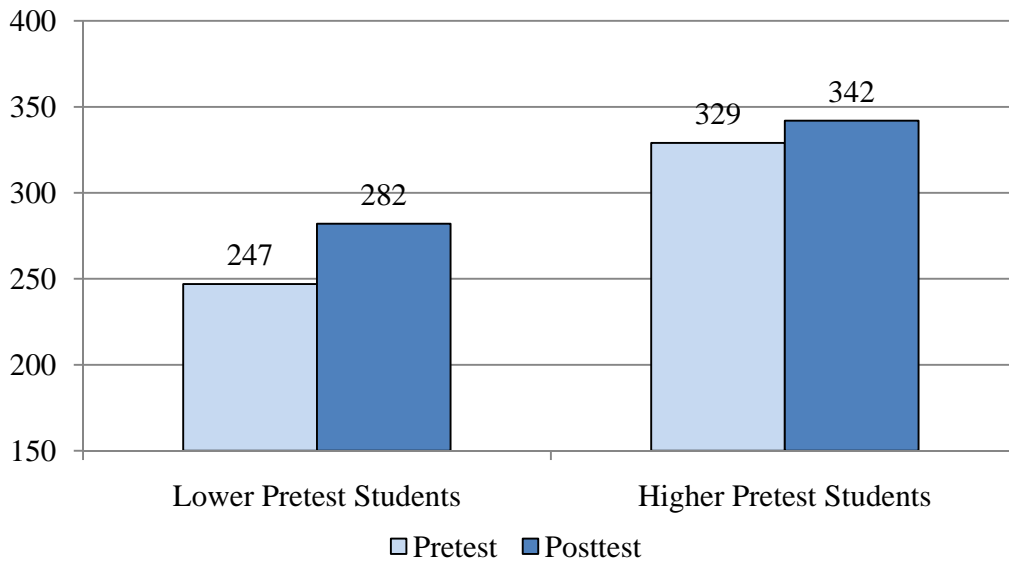
Pretest-to-posttest standard score comparisons are shown in Table 10 for the lower and higher Pretest scoring students. For both the higher and the lower scoring group, the total gain was statistically significant ($\leq .0001$). The effect size for the lower scoring group was large and for the higher scoring group the effect size was medium. However, the lower Pretest group increased 35 standard score points while the higher Pretest scoring group increased 13 standard score points.

Table 10
Grade 10 Paired Comparison t-test Results for Pretest/Posttest Standard Scores
for the High- and Low-Scoring Pretest Groups

Test	Test Form	Number of Students	Mean Standard Score	SD	t-test	Significance	Effect Size
Lower Scoring Group							
Total	Pretest	48	247.3	41.0	5.815	≤.0001	.90
Total	Posttest	48	281.6	36.7			
Higher Scoring Group							
Total	Pretest	48	329.1	23.2	4.679	≤.0001	.52
Total	Posttest	48	342.0	27.5			

Figure 8 provides a Pretest-to-posttest comparison of the total test standard scores of lower and higher scoring Pretest students.

Figure 8
Grade 10 Standard Score Increases for Lower and Higher Pretest Score Students



Conclusions

This study sought to determine the effectiveness of **Write Source 2012**®, an elementary and secondary writing program published by Houghton Mifflin Harcourt. The study was carried out with classes at grades 2, 4, 7 and 10. The teachers were using the program for the first time and received no special instruction in using the program.

Two research questions guided the study:

1. Is **Write Source 2012** ©effective in improving the writing skill strategies of elementary and secondary grade level students?
2. Is **Write Source 2012** effective in improving the writing skills and strategies of lower performing as well as higher performing elementary and secondary grade level students?

Question 1: Is Write Source 2012® an effective program in increasing the writing skills and strategies of both elementary and secondary grades?

A test which included both writing skills and strategy questions was developed to assess students at the beginning and end of a one-semester tryout of the program. Statistical analyses of students' scores at all grade levels showed that the students increased their scores statistically significantly on the Skills and Strategies test, the Writing assessment and the Total test scores. The effect sizes for the program for two of the comparisons were large, for 7 of the comparisons effect sizes were medium and for three comparisons the effect sizes were small.

Question 2: Is Write Source 2012® equally effective in increasing the writing skills and strategies of lower achieving students as well as higher achieving elementary and secondary students?

Statistical analyses of higher and lower pretest scoring students' scores at all grade levels showed that the students increased their scores statistically significantly on the Total test scores. The effect sizes for the program for five of the comparisons were large, for two of the comparisons effect sizes were medium and for one comparison the effect size was small.

Based on testing with both a test of writing skills and strategies as well as a writing assessment at grades 2, 4, 7 and 10, the conclusion is clear that students using the **Write Source 2012**® program made significant gains over the course of a one semester use of the program. The gains were consistent across all four grade levels. In addition, analysis of the **Write Source 2012**® student data clearly demonstrated that the program is effective with lower Pretest achieving students and higher achieving students.

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

- *The **Write Source 2012**® program is effective in increasing students' writing skills and strategies.*
- *The **Write Source 2012**® program is equally effective with lower Pretest scoring students and higher Pretest scoring students.*