A STUDY OF THE INSTRUCTIONAL EFFECTIVENESS OF
Integrated Mathematics © 2015
Report Number 509

August 2015

Advisory Board:

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

Jennifer M. Conner, Assistant ProfessorIndiana
University

Keith Cruse, Former Managing DirectorTexas
Assessment Program

Contents

Abstract

To help secondary school students develop the mathematics skills and depth of knowledge by integrating algebra and geometry concepts, along with data and statistics *Houghton Mifflin Harcourt* has published, *Houghton Mifflin Harcourt Integrated Mathematics © 2015* for students in grades 9 to 12*.*

In order to evaluate the program's effectiveness, *Houghton Mifflin Harcourt* contracted with the *Educational Research Institute of America* (ERIA) to conduct a one semester study to test the effectiveness of the program. The study was conducted with secondary students in grades 9 to 12 during the second semester of the 2014/2015 academic year. A small number of grade 8 students also participated in the study.

Pretest and post-test assessments were developed to cover the standards and content which the teachers agreed to complete during the second semester of the school year. Three teachers in one school participated in the study. Each of the three teachers used a different level of the program. The level 1 classes included students in grades 8 to 10. The level 2 and level 3 teachers included higher grade level students with level 3 being composed of mostly grade 11 and grade 12 students.

The teachers used the program as the primary component of their mathematics instruction. Teacher surveys showed the teachers used the program for an average of 4.5 days per week and about 40 minutes per class.

 The results showed that the *Houghton Mifflin Harcourt Integrated Mathematics* classes made statistically significant gains for all three levels of the program over the course of the one semester study. The effect size for the total group of students was large.

The total group was divided into two approximately equal groups of high pretest scoring students and low pretest scoring students. Analyses showed that both of these groups made statistically significant growth with a large effect size for the low pretest scoring group and a medium effect size for the high pretest scoring group. In fact, the lower pretest scoring group increase their test scores twice the increase of the higher pretest scoring group.

An independent analysis of each of the three level groups revealed that all three groups increased their scores statistically significantly. The effect size for level 1 was medium and for levels 2 and 3, the effect sizes were large.

Overview of the Study

This report describes a 2015 one-semester with students in grades 9 to 12 to determine the impact of the **Houghton Mifflin Harcourt Integrated Mathematics © 2015** program for secondary school students.

Houghton Mifflin Harcourt School Publishers contracted with the *Educational Research Institute of America* (ERIA) to conduct a one semester study during the second half of the 2014/2015 academic year study to determine the program's effectiveness. The **Houghton Mifflin Harcourt Integrated Mathematics © 2015** was the primary instructional program in the tryout classes.

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

Houghton Mifflin Harcourt Integrated Mathematics 1, 2, 3 offers an engaging and interactive approach to covering new state standards. This truly innovative high school mathematics program is designed for today's students with seamless integration of digital features for in-class and on-the-go learning. Focused, balanced, and rigorous instruction encompasses the philosophy and intent of the new state standards.

Houghton Mifflin Harcourt's Integrated Mathematics program for secondary students:

- *Helps prepare students for the Smarter Balanced Assessments through Houghton Mifflin Harcourt's Personal Math Trainer[®], powered by Knewton™, a digital adaptive assessment, intervention, and acceleration system that provides learning tools and real-time feedback to students, and reports student and classroom achievement to instructors*
- *Allows students to interactively explore new concepts through the Explore Activities, virtual manipulatives, Animated Math tutorials, video tutorials, and Personal Math Trainer's scaffolded assessment support*
- *Reinforces the Common Core Standards with Teacher Editions that provide Professional Learning with a focus on teaching for depth, and strategies for incorporating the Mathematical Practices into every lesson, as well as Professional Development videos.*
- *Delivers the next generation interactive digital Student Edition that works on all Internet-enabled devices—including tablets and smartphones—offers write-in functionality and note-taking capabilities, and contains point-of-use links to Math On the Spot Video Tutorials, Animated Math activities and simulations, and Personal Math Trainer online practice and help.*

Research Questions

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

1. Is *Houghton Mifflin Harcourt Integrated Mathematics* effective in helping secondary school students develop the mathematics skills and depth of knowledge by integrating algebra and geometry concepts, along with data and statistics in each course?

2. Is *Houghton Mifflin Harcourt Integrated Mathematics* effective in helping higher pretest scoring and lower pretest scoring students develop the mathematics skills and depth of knowledge by integrating algebra and geometry concepts, along with data and statistics in each course?

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 2014/2015 academic year in one school. The Level 1 and Level 2 programs included students in grades 8, 9, and 10. The Level 3 program included students in grades 10, 11, and 12. Each level of the program was taught by a different teacher.

Pre-tests and post-tests were administered at the beginning and end of the semester. The test were based on the assessments developed for the Integrated Mathematics program. Each of the three teachers indicated the program modules they would be teaching during the semester. Those modules were used as the basis for each of the three tests. The instructional modules included the following:

Level 1: Modules 7 to 13
Level 2: Modules 9 to 17
Level 3: Modules 9 to 13

The pretests and post-tests carefully matched the standards that were the focus of the identified instructional modules. Pretest and post-test administration was under the direction of the classroom teacher. All tests were returned to ERIA for scoring and analysis.

Timeline and Program Use

The teachers used the *Houghton Mifflin Harcourt Integrated Mathematics* text as their primary instructional program. The teachers reported using the program an average of 4.5 days per week and for an average of about 40-45 minutes per day over the entire semester. Pretests were administered the beginning of January, 2015 and posttests were administered the end of May, 2015.

Description of the Research Sample

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

Table 1
Demographic Characteristics of the School

School	State	Location	Grades	Enrollment	% Minority	% Free/Reduced Lunch
1	WI	Rural	9-12	1440	22%	42%

Description of the Assessments

The pretest and posttest used in the study were developed to assess standards-based on the literary analysis of various texts. Based on these standards 40 item multiple-choice assessment pre/post tests were developed focusing on mathematics skills and depth of knowledge by integrating algebra and geometry concepts, along with data and statistics.

Table 2 provides the statistical results for the administration of the pretest and the post-test for all three Integrated Math pretests and post-tests and the averages for the pretests and post-tests. The KR 20 Reliability and the Standard Error of Measurement for the post-test 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

Test	Reliability*	SEM**
Integrated Math 1 Pretest	.59	2.68
Integrated Math 1 Post-test	.68	2.56
Integrated Math 2 Pretest	.84	2.95
Integrated Math 2 Post-test	.79	2.28
Integrated Math 3 Pretest	.63	2.88
Integrated Math 3 Post-test	.63	2.88
Average Pretests	.69	2.85
Average Post-tests	.70	2.73

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 *Integrated Mathematics* assessments. 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:

- For the total group of students across all 3 levels of the program a paired comparison *t*-test was used to compare the pretest mean standard scores with the posttest mean standard scores for all students. Standard score comparisons were used to equate the scores across the three levels.
- Descriptive statistics were also used to compare pretest and post-test standard test scores for the total group as well as the higher and lower pretest score groups.
- The total group students was 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 program was equally effective with students who had lower and higher pretest scores. As with the total group standard scores were used for the comparisons.
- For each of the three groups, levels 1, 2, and 3, paired comparison *t*-tests were computed for each group.

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

Across All Three Levels

A paired comparison *t*-test to determine if the difference from pretest standard scores to posttest standard scores for the total group of students was statistically significant. For this analysis, pretest and posttest scores were matched for a total of 115 students. Students who did not take both the pretest and the posttest were not included in the analysis.

Table 3 shows that the average standard score on the pretest was 281, and the average standard score on the posttest was 319. The increase was statistically significant ($\leq$.0001). The effect size was large.

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

Test	Number Students	Mean Standard Score	SD	t-test	Significance	Effect Size
Pretest	115	281	46.0	8.381	$\leq$.0001	.83
Posttest	115	319	46.1			

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 115 students was divided into two approximately equal sized groups of 57 for the lower scoring students and 58 for the higher scoring students. The lower scoring had a mean score of 247 with scores ranging from 179 to 313. The higher scoring group scored an average standard score on the pretest of 314 with scores ranging from 276 to 383.

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 ($\leq$.0001). The effect size for the lower pretest scoring group was large and for the higher pretest scoring the effect size was medium. In line with those results, the data shows that the lower pretest group increased 58 standard score points and the higher pretest scoring group increased 20 standard score points.

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

Test Form	Number Students	Standard Score	SD	t-test	Significance	Effect Size
Lower Scoring Group						
Pretest	57	247	33.0	8.269	≤.0001	1.35
Posttest	57	305	51.0			
Higher Scoring Group						
Pretest	58	314	29.5	4.014	≤.0001	.62
Posttest	58	334	35.6			

Figure 1 provides a pretest-to-posttest comparison of the standard scores of the total group and the 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 67 point difference at the beginning of the academic year and a 29 point difference by the end of the academic year.

Figure 1
Standard Score Increases for Students and for Subgroups of
Lower and Higher Pretest Scoring Students
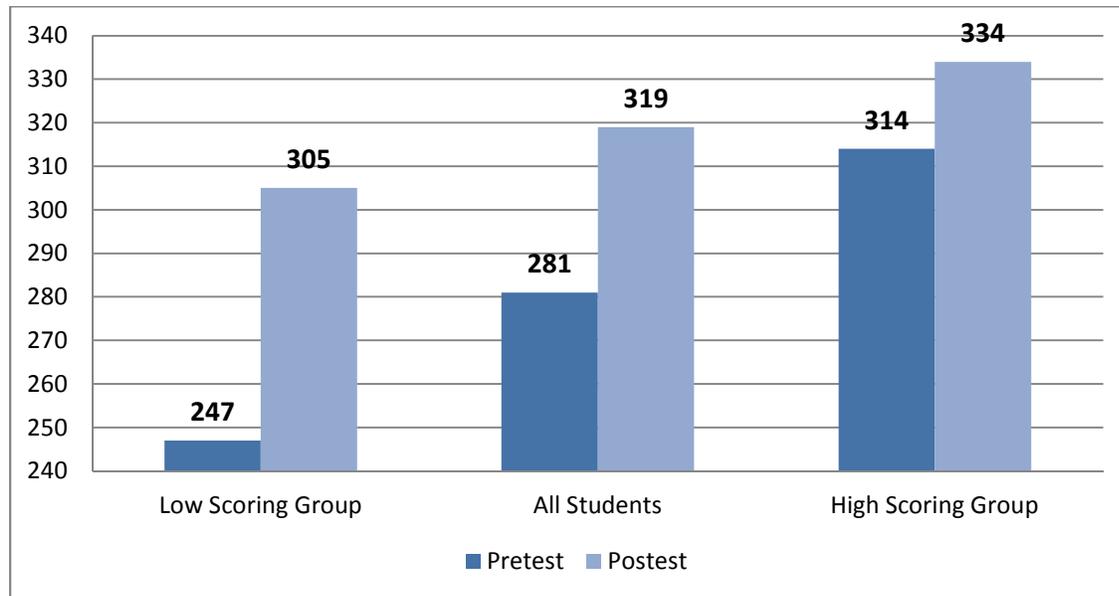
Analyses for Each Program Level

ERIA conducted a paired comparison *t*-test to determine if the difference from pretest standard scores to posttest standard scores was statistically significant for each of the program level groups. These three analyses were somewhat problematic because the number of students in each group was small. For Level 1, the pretest and posttest standard were available for 51 students. For Level 2 the number of students was 41, and for Level 3, the total number of students was 23. Students who did not take both the pretest and the posttest were not included.

Table 5 shows that the average standard score increase averaged about 40 standard score points. The increase was statistically significant for all three programs. The effect size was medium for the Level 1 program and large for the Level 2 and 3 programs.

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

Test	Number Students	Mean Standard Score	SD	t-test	Significance	Effect Size
Level 1 Program						
Pretest	51	282	43.9	5.223	≤.0001	.77
Post-test	51	318	49.8			
Level 2 Program						
Pretest	41	279	47.7	5.637	≤.0001	.92
Post-test	41	321	43.3			
Level 3 Program						
Pretest	23	281	49.2	3.369	≤.003	.81
Post-test	23	319	44.1			

Conclusions

This study sought to determine the effectiveness of *Houghton Mifflin Harcourt Integrated Mathematics © 2015,* a grade 9 to 12 mathematics program published by Houghton Mifflin Harcourt. The study was carried out with secondary school classes and included program levels 1, 2, and 3.

Each program level was taught by a different teacher. Classes included students from grade 8 to grade 12. The teachers were using the program for the first time and received no special instruction in using the program. Teachers used the program for an average of 4.5 days per week and for about 40 minutes for each class.

Two research questions guided the study:

*Question 1: Is **Houghton Mifflin Harcourt Integrated Mathematics** effective in helping secondary school students develop the mathematics skills and depth of knowledge by integrating algebra and geometry concepts, along with data and statistics in each course?*

Pretests and post-tests were developed to match the standards of the Integrate Mathematics program for Levels 1, 2, and 3. The assessments covered only those units that teachers were expected to complete during the second semester of the 2014-2015 academic year. For the total group of students across the combined group of students for each the three levels, statistical analyses of students' scores showed that the students increased their scores statistically significantly and the effect size was large.

The sample size for each of the three program level groups was small. However, the results comparing pretest to post-test scores produced statistical significance for each of the three programs. The effect size for the Level 1 program was medium while the effect sizes for both the Level 2 and Level 3 programs were large.

*Question 2: Is **Houghton Mifflin Harcourt Integrated Mathematics** effective in helping higher pretest scoring and lower pretest scoring students develop the mathematics skills and depth of knowledge by integrating algebra and geometry concepts, along with data and statistics in each course?*

The total group of all 115 students across all three levels of the program was sorted on the basis of each students pretest score. The students who scored lowest on the pretest were considered the lower performing group. While those students who scored highest on the pretests were considered the high scoring group.

A dependent sample *t*-test was used to determine differences from pretesting to post-testing. The results showed that for both groups the increases were statistically significant. For the lower scoring group the effect size was large and for the higher scoring group the effect size was medium. A comparison of the standard score point increases showed that the lower scoring group increased their average score by more than twice the average point increase of the higher pretest scoring group.

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

- The ***Houghton Mifflin Harcourt Integrated Mathematics*** is effective in helping secondary school students develop the mathematics skills and depth of knowledge by integrating algebra and geometry concepts, along with data and statistics in each course?

- The ***Houghton Mifflin Harcourt Integrated Mathematics*** is effective in helping higher pretest scoring and lower pretest scoring students develop the mathematics skills and depth of knowledge by integrating algebra and geometry concepts, along with data and statistics in each course?