Assessing the Effects of *McGraw-Hill's My Math* Program On Student Achievement and Teacher Attitudes: Final Report

Shannon Flaum-Horvath, PhD

November, 2012

SKF EDUCATIONAL SERVICES, LLC

RESEARCH AND EVALUATION CONSULTANTS

Table of Contents

Executive Summary

Introduction

In the fall of the 2011 school year, *McGraw-Hill*, in conjunction with *SKF Educational Services*, conducted a year-long study investigating the effects of *McGraw-Hill My Math* on student achievement and teacher attitudes. The study utilized a pretest-posttest, randomized control-group design, with assignment at the classroom level. The sample contained 829 students in Kindergarten, second, and fifth-grade; 494 students comprised the *McGraw-Hill My Math* condition, and 335 comprised the control condition. To assess the effect of *McGraw-Hill My Math* on student achievement, students in the *McGraw-Hill My Math* and the control conditions were pretested and post-tested using the math subject test of the *Terra Nova*. Qualitative measures, including surveys, interviews, and classroom observations, were utilized to determine fidelity of program implementation and capture teacher reactions to core program components.

The purpose of this study was to determine how *McGraw-Hill My Math* currently 'worked', in a naturalistic setting. Specific questions of interest include:

1. What effect does *McGraw-Hill My Math* have on students' academic achievement, as measured by performance on the math portion of the *Terra Nova*?
2. How does the performance of students using the *McGraw-Hill My Math* program compare to the performance of students not using *McGraw-Hill My Math*?
3. What effect do student demographic variables have on students' performance?
4. What effect does *McGraw-Hill My Math* have on the attitudes and behaviors of students and educators who are exposed to the program?
5. To what degree is *McGraw-Hill My Math* implemented, in a natural setting?
6. What recommendations, if any, do teachers and students provide for improving the program?

Teachers did not receive training, as the program was under development. Print materials were made available in October of 2011, and access codes for digital resources became available in November of 2011. The online components were being added to the digital platform throughout the entire course of the study. Since many of the *McGraw-Hill My Math's* lessons rely on the availability of the digital resources, not having access to them may have impacted the overall effectiveness of the study.

Key Findings

Demographics and Location of Sites

- Across all grade levels, the *McGraw-Hill My Math* group had a higher percentage of males (50%) than did the control group (roughly 47%).
- The ethnic distribution revealed similar percentages of students identified as African American, Caucasian, and multi-ethnic; however, there was a higher percentage of Hispanic students and a lower percentage of Asian students in the control group.
- Groups were fairly equivalent in the percentage of ESL and special needs students.
- Across all grade levels, 64% of the *McGraw-Hill My Math* classrooms had a free/reduced lunch population greater than 50%, while 56% of control classrooms had a free/reduced lunch population greater than 50%.
- Teachers in the *McGraw-Hill My Math* and control groups were similar in terms of years teaching experience.
- Sites included classrooms from various geographic areas and types of community.

Student Achievement[1]

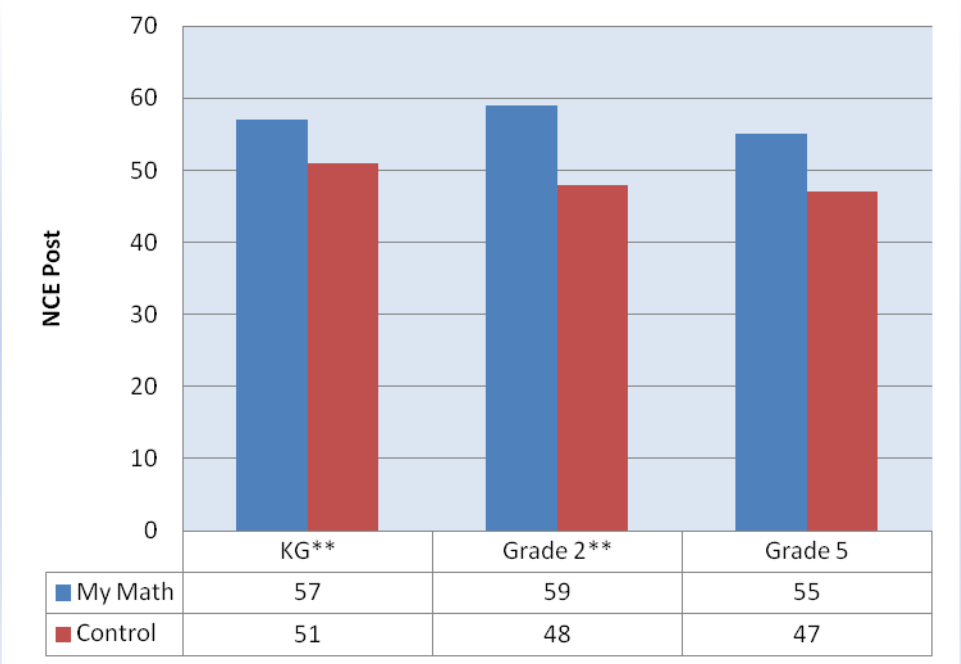
- Kindergarten. The average posttest scaled score on the *Terra Nova* was 495.24 for the *McGraw-Hill My Math* group, compared to 484.85 for the control group. After controlling for differences on the pretest and demographic variables, this difference was ***not*** found to be statistically significant. The average posttest normal curve equivalent score was approximately 57 for the *McGraw-Hill My Math* group and 51 for the control group. The difference in normal curve equivalent posttest scores ***was*** found to be statistically significant, with an average increase of 6 normal curve equivalent units.

- Second Grade. The average posttest scaled score on the *Terra Nova* was 592.73 for the *McGraw-Hill My Math* group, compared to 571.21 for the control group. The average normal curve equivalent posttest score on the *Terra Nova* was 59 for the *McGraw-Hill My Math* group and 48 for the control group. After controlling for differences on the pretest, the difference ***was*** found to be statistically significant, for ***both*** the scaled score and the normal curve equivalent score. For students in the *McGraw-Hill My Math* group, the gain was approximately 14 scaled score units and 6 normal curve equivalent units.

[1] See the Discussion (page 32) for further insight into the interpretation of the data

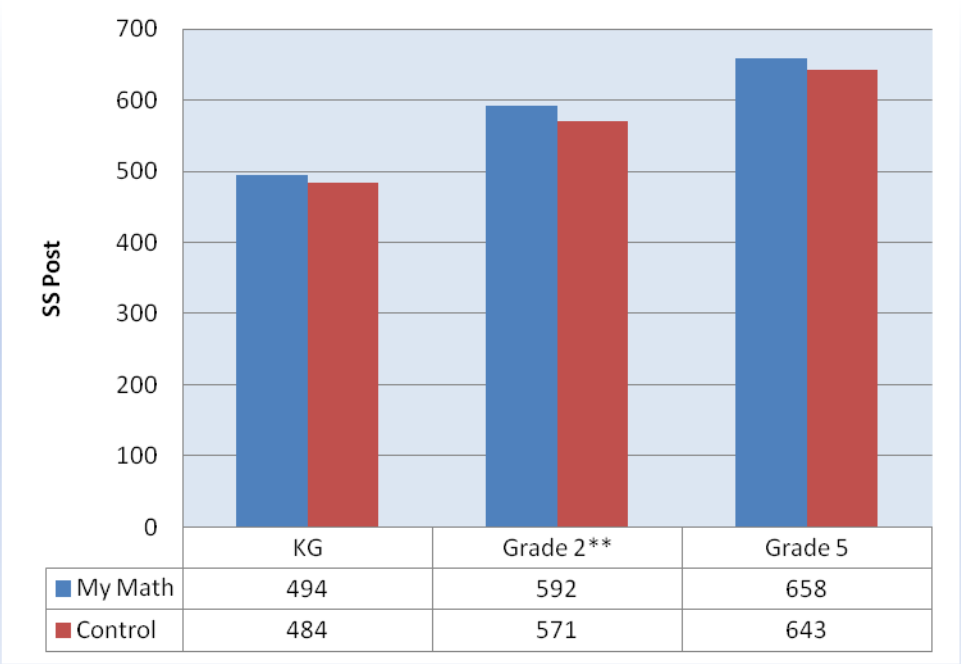
- Fifth Grade. The average posttest scaled score on the *Terra Nova* was 658.51 for the *McGraw-Hill My Math* group, compared to 643.14 for the control group. The average normal curve equivalent posttest score on the *Terra Nova* was approximately 55 for the *McGraw-Hill My Math* group and 47 for the control group. After controlling for differences on the pretest, the difference was ***not*** found to be statistically significant, for either the scaled score or the normal curve equivalent score.

Comparison of McGraw-Hill My Math and Control Students, NCE posttest



**Results statistically significant, $p < .05$

Comparison of My Math and Control Students, SS posttest



	KG	Grade 2**	Grade 5
My Math	494	592	658
Control	484	571	643

**Results statistically significant, $p < .05$

Fidelity of Implementation

- With the exception of one teacher (not included in the analysis), all teachers utilized the *Student Interactive Text*, the primary component of the program.
- Eight teachers, (36%) reported that they accessed materials from other sources immediately prior to conducting state-mandated assessments.
- About 25% of teachers reported that they did not use the *Vocabulary Cards* and the *Foldables* on a daily basis.
- 60% of teachers reported that they did not access the online features on a regular basis.

Teacher Attitudes

- 75% of teachers reported they would recommend *McGraw-Hill My Math* to a colleague.
- 90% of teachers reported that *McGraw-Hill My Math* is of high instructional quality.
- 90% of teachers reported that *McGraw-Hill My Math* reflects the breadth and depth of the *Common Core Standards*.
- 85% of teachers reported that *McGraw-Hill My Math* supports 21st Century skills.
- 95% of teachers reported that *McGraw-Hill My Math* instructional plan and lesson preparation was manageable.
- 95% of teachers reported that their students felt successful during *McGraw-Hill My Math* instruction.
- 95% of teachers reported that their students were fully engaged during *McGraw-Hill My Math* instruction.
- Descriptions of the program include the terms, "colorful", "engaging", and "fun" (72% of responses); "easy to use", "teacher-friendly", "student-friendly" (55% of responses), and "standards-based" (40% of responses).

Introduction

"The eminence, safety, and well-being of nations have been entwined for centuries with the ability of their people to deal with sophisticated quantitative ideas. Leading societies have commanded mathematical skills that have brought them advantages in medicine and health, in technology and commerce, in navigation and exploration, in defense and finance, and in the ability to understand past failures and to forecast future developments"

– Foundations for Success: The Final Report of the National Mathematics Advisory Panel, 2008, p. 1

The Problem

Findings from the 2011 National Assessment of Education Progress (NAEP) reveal that although student proficiency is inching upward, U.S. students generally continue to exhibit lower than desired levels of math proficiency. While the percentages of fourth-grade students scoring at or above the *Proficient* and *Advanced* levels on the NAEP were higher in 2011 than in previous assessment years, the percentage of students scoring at or above *Basic* did not change significantly from 2009 to 2011. Further, there remain disparities in mathematics achievement related to ethnicity, income, disability status, and language proficiency: of the lowest performing students, 62% were identified as African American or Hispanic, 74% were eligible for free/reduced lunch, 24% were identified with learning disabilities, and 22% were identified as English Language Learners (National Center for Education Statistics, 2011, p. 11).

The 2007 *Trends in International Mathematics and Science Study* (TIMSS) provides an international comparison of the performance of U.S. students relative to that of their peers. Of the 35 countries participating in the TIMSS, the average fourth-grade U.S. mathematics score was higher than those in 23 countries, lower than those in 8 countries (all of which are in Asia or Europe), and not measurably different from those in 4 countries (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008, p. 6).

While it appears that our students are making gains, somewhere along the way many students failed to develop the foundational skills necessary to achieve in mathematics. As many states move toward adopting the *Common Core State Standards*, it becomes imperative that curricula incorporate content in a way addresses the standards while maximizing student achievement.

Overview of the Study

This report provides a detailed examination of the procedure, methods and findings from the 2011-2012 implementation of *McGraw-Hill's My Math* program. Initiated in the fall of the 2011-2012 school year, the study utilizes a pretest-posttest, randomized control-group design, with assignment implemented at the classroom level. The sample contains students and teachers in targeted grades (Kindergarten, 2nd, and 5th), distributed across multiple states and types of community. Included in the study design are quantitative and qualitative measures intended to capture the effects of *McGraw-Hill My Math* on student achievement and teacher reactions to core program components.

It should be noted that this study is both formative and summative in nature. Information gathered during this study will help provide valuable information regarding the program's efficacy and the recommendations of the participants will guide further resource development as the program is being fine-tuned for full release.

The McGraw-Hill My Math Program

In effort to bolster student achievement in mathematics, *McGraw-Hill* has published *McGraw-Hill My Math* a unique, research-based math program currently in development. Emphasizing vocabulary, problem solving, and real-world applications, *McGraw-Hill My Math* addresses the 'big concepts' included in the *Common Core State Standards*. Designed to foster student engagement, *McGraw-Hill My Math* includes a high-quality, visually-appealing interactive textbook that helps students take ownership over their learning to make it truly their own. The program blends together a unique print component with highly interactive digital tools to provide a total, comprehensive solution for elementary mathematics instruction. There needs to be some consideration in the final analysis of this study of the lag in the development of the digital resources behind those of print where the result could have hampered the student experience.

Purpose of Study

The purpose of this study was to determine how *McGraw-Hill My Math* currently 'works' in the classroom setting and to gather suggestions to make it more effective prior to its full release. Specific questions of interest include:

7. What effect does *McGraw-Hill My Math* have on students' academic achievement, as measured by performance on the math portion of the *Terra Nova*?
8. How does the performance of students using the *McGraw-Hill My Math* program compare to the performance of students not using the *My Math* program?
9. What effect do student demographic variables have on students' performance?
10. What effect does *McGraw-Hill My Math* have on the attitudes and behaviors of students and educators who are exposed to the program?
11. To what degree is *McGraw-Hill My Math* implemented, in a natural setting without direct support or program training?
12. What recommendations, if any, do teachers and students provide for improving the program?

Post-hoc analysis may also be included. In an effort to address these questions, this study incorporates quantitative and qualitative information obtained through assessments, surveys, and interviews.

Method

Selection of Sites

In late August/early September 2011, the selection of sites and recruiting of teachers was initiated by *Saperstein Associates*, a public opinion research company located in Columbus, Ohio. At the start, 40 classrooms incorporating students in Kindergarten, 2nd, and 5th grades were targeted for the study. Email blasts, containing pertinent information about the study, were sent to teachers from around the country. Those teachers expressing interest were administered a short survey to screen out participants through use of exclusionary criteria (e.g., current use of a *McGraw-Hill* mathematics program and a lack of access to computers) and to collect teacher and school demographic variables. To compensate teachers for their time and efforts, it was explained that a $250 cash honorarium was to be provided, contingent upon completing study requirements. Teachers, as well as administrators, wishing to participate in the study were asked to provide written consent. After screening out ineligible participants and upon receipt of consent forms, *Saperstein Associates* forwarded all information to a research consultant at *SKF Educational Services, LLC* who contacted teachers and informed them of their primary responsibilities and general layout of the study. Teachers were informed from the start that a requirement for participation included their consent to random assignment to either the *McGraw-Hill My Math* group or the control group. Those in the *McGraw-Hill My Math* group would receive all print materials as well as access to the digital features of the program as they became available. Those teachers selected for the control group were asked to continue to use their existing math program. While several teachers initially expressed reluctance about consenting to random assignment, once the study design was further explained and teachers understood the importance of the assignment, all teachers provided consent. With recruiting completed, classrooms were stratified into groups according to geographic location (West, Midwest, South, East) and type of community (suburban, rural, urban). At that time, teachers were randomly assigned to teach either *McGraw-Hill My Math* or to continue to use their current math program.

It is worthy to note two challenges that presented themselves at the start. First, recruiting of participants coincided with the start of the school year for many districts, understandably a very busy time for teachers. It is possible that teachers, who may have participated if given more planning time, could not commit to a full year of implementation and were reluctant to sign on. Second, recruitment efforts stalled at the beginning of September, as Hurricane Irene ventured up the east coast and caused massive power outages. While assignment of teachers to respective groupings was delayed to provide teachers from the eastern part of the country the opportunity to respond, relatively few did so. As such, teachers from the eastern states are not adequately represented. Beyond this, it is not known with certainty how or to what degree these circumstances may have impacted the nature of the participant pool.

Description of Research Sites and Participants

The sites selected for this study initially included 40 classrooms contained within 29 different public school districts and 6 states (CO, IL, MD, NC, NY, OH). Schools were stratified by geographic location and community, with an even representation of schools located within rural, suburban, and urban areas. Shortly after program inception, one participating teacher found it necessary to withdraw from the study after she received a new teaching assignment, and one teacher found it necessary to withdraw for personal/health reasons. It is worthy to note that both teachers were initially assigned to the control condition. This rendered 38 classrooms and a total of 829 students available for review. Twenty-two classrooms and 494 students comprise the *McGraw-Hill My Math* condition, and 16 classrooms and 335 students comprise the control condition[1].

Description of Sites

Classrooms were stratified according to geographic area and type of community (see Table 1.). Five *McGraw-Hill My Math* classrooms are located within the Western part of the Unites States, ten are located within the Midwest, five are located within the South, and two are located in the East. Three control classrooms are located in the West, eight are located in the Midwest, and five are located in the South. No control classrooms are located in the East.

[1]Deliberate over assignment to those in the *McGraw-Hill My Math* group to allow for attrition.

Table 1

Distribution of Classrooms, by Geographic Area and Community

		*My Math**	*Control**
Geographic Area	***Community***		
West (CO)	Urban	1	0
	Suburban	4	2
	Rural	0	1
Midwest (IL, OH)	Urban	1	2
	Suburban	4	3
	Rural	5	3
South (NC)	Urban	1	2
	Suburban	1	1
	Rural	3	2
East (MD, NY)	Urban	0	0
	Suburban	1	0
	Rural	1	0

*Note: Percentages are calculated based on number of classrooms per grouping (*My Math* or Control).

Description of Teachers

As part of the recruitment process, teachers were asked to provide information about their experience and education in teaching (see Table 2). Teachers assigned to the *McGraw-Hill My Math* condition, have on average 10.3 years of teaching experience, with an average of 9.9 years teaching math specifically. Teachers assigned to the control condition have on average 11.3 years of teaching experience, with an average of 10.4 years teaching math, specifically. The difference between the groups in years of teaching experience is not statistically significant, $t(37) = -.453$, $p = .654$. Similarly, the difference between the groups in years of teaching math is not statistically significant, $t(37) = -.863$, $p = .394$. Teachers' level of education is roughly equivalent, with about 40% of teachers possessing a Bachelor's degree and 60% of teachers possessing a Master's degree.

Table 2

Comparison of My Math and Control Groups, Selected Teacher Variables

		My Math	*Control*
Years Teaching			
	5 or less	27% (n=6)	19% (n=3)
	6 to 15	36% (n=8)	38% (n=9)
	15 or more	32% (n=7)	31% (n=5)
Years Teaching Math			
	5 or less	27% (n=6)	19% (n=3)
	6 to 15	38% (n=8)	50% (n=8)
	15 or more	27% (n=6)	31% (n=5)
Education Level			
	Bachelors	41% (n=9)	38% (n=6)
	Masters	59% (n=13)	62% (n=10)

Classroom Variables

Table 3. provides an analysis of the number of *McGraw-Hill My Math* and control classrooms by percentage of ESL and Free/Reduced lunch population. In both groups, the majority of classrooms were represented by lower percentages of ESL/ELL students and higher percentages of students receiving free or reduced price meals. The number of classrooms with low (10% or less), medium (11 to 50%) and high (greater than 50%) percentages of students characterized as ESL was fairly consistent for the *McGraw-Hill My Math* and control classrooms. About 64% of the *McGraw-Hill My Math* classrooms had a free/reduced lunch population greater than 50%, while 56% of control classrooms had a free/reduced lunch population greater than 50%.

Table 3

Comparison of My Math and Control Groups, Selected Classroom Variables

		My Math	***Control***
ESL or ELL			
	10% or less	72% (n=16)	81% (n=13)
	11 to 50%	14% (n=3)	13% (n=2)
	>50%	9% (n=2)	6% (n=1)
Free/Reduced Lunch			
	10% or less	5% (n=1)	0% (n=0)
	11 to 50%	32% (n=7)	44% (n=7)
	>50%	64% (n=14)	56% (n=9)

Description of Students

A comparison of the *McGraw-Hill My Math* and control groups on selected student characteristics is presented in Table 4. Across all grade levels, the *McGraw-Hill My Math* group had a higher percentage of males (50%) than did the control group (roughly 47%). The ethnic distribution revealed similar percentages of students identified as African American, Caucasian, and multi-ethnic; however, there was a higher percentage of Hispanic students and a lower percentage of Asian students in the control group. A higher percentage of students reported as receiving free or reduced lunch were found in the control group; however, the status of 25% of students in the *McGraw-Hill My Math* group and 18% of those in the control group was not released by corresponding districts. Groups were fairly equal in terms of percentage of English as a Second Language (ESL) population, and were equivalent in the percentage of students receiving special needs services.

Table 4

Comparison of My Math and Control Groups, Selected Student Variables

	My Math (n = 477)	Control (n = 354)
Characteristic		
Gender		
Female	238 (49.9)*	185 (52.3)
Male	239 (50.1)	168 (47.5)
Grade		
Kindergarten	151 (31.7)	103 (29.1)
2nd	155 (32.5)	115 (32.5)
5th	171 (35.8)	136 (38.4)
Ethnicity		
African American	93 (19.5)	74 (20.9)
Caucasian	291 (61.0)	214 (60.5)
Hispanic	42 (8.8)	50 (14.1)
Asian	15 (3.1)	3 (0.8)
Multi-Ethnic	3 (2.7)	9 (2.5)
Other Reported	3 (0.6)	3 (0.8)
Free/Reduced Lunch		
Yes	165 (34.6)	144 (40.7)
No	197 (41.3)	146 (41.2)
Unknown	115 (24.1)	64 (18.1)
ESL/ELL		
Yes	45 (9.4)	35 (9.9)
No	432 (90.6)	318 (89.8)
Special Education		
Yes	43 (9.0)	32 (9.0)
No	434 (91.0)	322 (91.0)

*Note: values indicate number of students, followed by percentage in parentheses.

Program Implementation

While it was the desire of the research team to implement *McGraw-Hill My Math* at the start of the school year, the abovementioned circumstances during the recruiting phase and a delay in the release of the *Terra Nova* assessment lead to implementation in late October. At that time, teachers received all print materials including teaching materials and student workbooks. As the digital resources were not yet fully operational, teachers did not immediately receive login codes for accessing technology. At the outset, the project team wished to determine the ease in which teachers were able to 'simply pick up and use' the program, as would be typical in a natural setting. For this reason, and since the program was in development during study inception, no formal training was conducted.

The ability to infer causality is critically tied to the degree to which a program is implemented with fidelity, and various methods were employed to fully determine how teachers were using the program. Specifically, the team wished to determine:

- Are teachers in fact using the *Student Interactive Text*?
- Are teachers referencing the *Teacher's Edition* when providing instruction?
- To what degree are teachers following the curriculum, as written?
- How are teachers utilizing the digital components?

Throughout the study, an independent research consultant provided ongoing support to the teachers, assisted with questions, and provided feedback, as needed. Weekly check-ins were conducted via email to determine that teachers were using the program as intended and to document changes that may have compromised the integrity of the program. Findings obtained during check-ins were triangulated with questions regarding implementation fidelity on the survey and during the interviews. While it was the desire of the team to conduct observations of every classroom, financial constraints and travel restrictions prevented this from occurring; however, classroom observations were conducted for a random sample of nine classrooms. Based on the information collected, intervention integrity was determine by combining the data obtained from the weekly check-ins, surveys, interviews, and classroom observations.

Teachers were asked during the weekly email check-in, on the survey, and during the interview, whether or not they were in fact using the *Student Interactive Text*. All but one teacher reported that he/she was using the *Student Interactive Text*. Classroom observations of selected teachers and information provided from the survey corroborated with this finding. All teachers, (with the exception of the sole teacher not using the *Student Interactive Text*) reported that they utilized the *Teacher's Edition* when providing instruction and this was in fact observed during the classroom observations. There was, however, some variation in terms of how closely the program was implemented as designed. Several teachers (n = 8) openly admitted that they

"pulled information from other sources" when preparing students for the upcoming state-mandated assessments. About 25% of teachers reported that they did not regularly use the *Vocabulary Cards* and the *Foldables*, and for reasons cited later in this report, over half (60%) of teachers did not regularly access the technological features of the program.

Research Design and Data Analysis

This study utilized a pretest-posttest, randomized control group experimental design, with students assigned to group at the classroom level. The math subtests of the *Terra Nova* (CTB/McGraw-Hill) served as the pretest and posttest. Students in the *McGraw-Hill My Math* and control groups were pretested in October of 2011 and posttested in May of 2012.

Student performance on the *Terra Nova* was considered in terms of normal curve equivalents and scaled scores.

Results

Student Achievement

Descriptive statistics and the results of tests for statistical significance are separately listed for Kindergarten, second-, and fifth-grade students.

Kindergarten

The average *Scaled Scores* (SS) on the pretest and posttest administration of the *Terra Nova*, and the average *Normal Curve Equivalent* (NCE) scores on the posttest, by group, are provided in Table 5. *Normal Curve Equivalent* (NCE) pretest scores are not provided, as there are no national norms generated for Kindergarten students in the fall of their Kindergarten year.

Table 5

Descriptive Statistics for SS and NCE Scores by Group, Kindergarten

Measure	Group	*M*	*Mdn*	*SD*	*Skew*	*Range*
SS pretest						
	My Math	454.29	455.50	38.62	0.30	[351, 629]
	Control	443.51	442.00	29.51	0.26	[290, 532]
SS posttest						
	My Math	495.24	497.00	66.08	-0.40	[290, 629]
	Control	484.85	486.00	36.35	0.40	[389, 629]
NCE posttest						
	My Math	57.60	57.50	26.39	-0.14	[13,99]
	Control	51.59	50.00	16.65	0.27	[1,85]

On average, students in the *McGraw-Hill My Math* group performed about 11 units higher on the *SSpretest* than did students in the control group. This finding held true for the *SSposttest*, as students in the *McGraw-Hill My Math* group outperformed students in the control group by 11 units. To control for differences in initial performance on the *Terra Nova* pretest, linear regression was utilized to determine the program effect on students' posttest scaled scores. After controlling for students' *SSpretest* score, the effect of program (unstandardized regression coefficient $b = 2.99$) was *not* found to be statistically significant, $t_{228} = .430$, $p = .668$. This

finding remained after incorporating gender, ethnicity, free/reduced lunch status, ESL status, and special education status into the statistical model.

To determine the significance of the difference in NCE posttest scores, an independent samples t-test was performed. Results indicate that the difference in NCE posttest scores between the *McGraw-Hill My Math* and control group *was* found to be statistically significant, $t_{228} = 2.138$, p = .000. On average, students in the *McGraw-Hill My Math* group outperformed students in the control group by approximately 6 NCE units. However, the 95% Confidence Interval is [.47 – 11.56], indicating that the likely range of increase may be as small as .47 NCE units or as large as 11.56 NCE units.

Second Grade

The average *Scaled Score* (SS) and *Normal Curve Equivalent* (NCE) scores on the pretest and posttest administration of the *Terra Nova*, by group is provided in Table 6. On average, those second grade students in the *McGraw-Hill My Math* group exhibited higher SS and corresponding NCE pretest scores than did second grade students in the control group, by 16 points and 6 points, respectively.

Table 6

Descriptive Statistics for Normal Curve Equivalent Scores by Group, Second Grade

Measure	Group	*M*	*Mdn*	*SD*	*Skew*	*Range*
SS pretest						
	My Math	554.55	552.00	36.95	0.29	[450, 676]
	Control	538.53	544.00	64.14	0.27	[450, 720]
SS posttest						
	My Math	592.73	592.00	34.23	0.27	[502, 720]
	Control	571.21	571.00	42.43	-0.37	[448, 664]
NCE pretest						
	My Math	55.99	54.00	18.62	0.19	[8, 99]
	Control	50.19	50.00	20.42	0.21	[1, 86]
NCE posttest						
	My Math	59.08	60.00	17.16	-0.37	[1, 99]
	Control	48.27	47.00	21.52	-0.17	[1, 99]

As with the Kindergarten data, linear regression was utilized to determine the program effect on students' posttest SS and NCE score. After controlling for students' pretest SS, program condition was found to be a statistically significant predictor of posttest SS (unstandardized regression coefficient $b = 13.87$, $t_{237} = 3.57$, $p = .000$). After controlling for students' pretest NCE score, the program condition was found to be a statistically significant predictor of posttest NCE scores, (unstandardized regression coefficient $b = 6.19$ $t_{237} = 4.165$, $p = .000$). These findings did not change, in either case, when gender, ethnicity, free/reduced lunch status, ESL/ELL status, or special education status were entered into the statistical model.

Fifth Grade

The average *Scaled Score* (SS) and *Normal Curve Equivalent* (NCE) scores on the pretest and posttest administration of the *Terra Nova*, by group is provided in Table 7. On average, those fifth grade students in the *McGraw-Hill My Math* group exhibited higher SS and corresponding NCE pretest scores than did students in the control group, by 15 points and 8 points, respectively.

Table 7

Descriptive Statistics for Normal Curve Equivalent Scores by Group, Fifth Grade

Measure	Group	*M*	*Mdn*	*SD*	*Skew*	*Range*
SS pretest						
	My Math	647.37	646	34.09	-0.13	[534, 717]
	Control	633.48	632	36.24	-1.77	[430, 698]
SS posttest						
	My Math	658.51	658	46.68	-1.69	[430, 750]
	Control	643.14	651	45.67	-2.05	[430, 748]
NCE pretest						
	My Math	54.70	53	19.63	0.30	[8, 99]
	Control	47.81	45	16.84	-.04	[1, 86]
NCE posttest						
	My Math	54.89	52	20.92	-0.14	[1, 99]
	Control	46.98	48	18.28	-0.09	[1, 99]

As with the Kindergarten and second-grade data, linear regression was utilized to determine the program effect on fifth grade students' SS and NCE scores. After controlling for students'

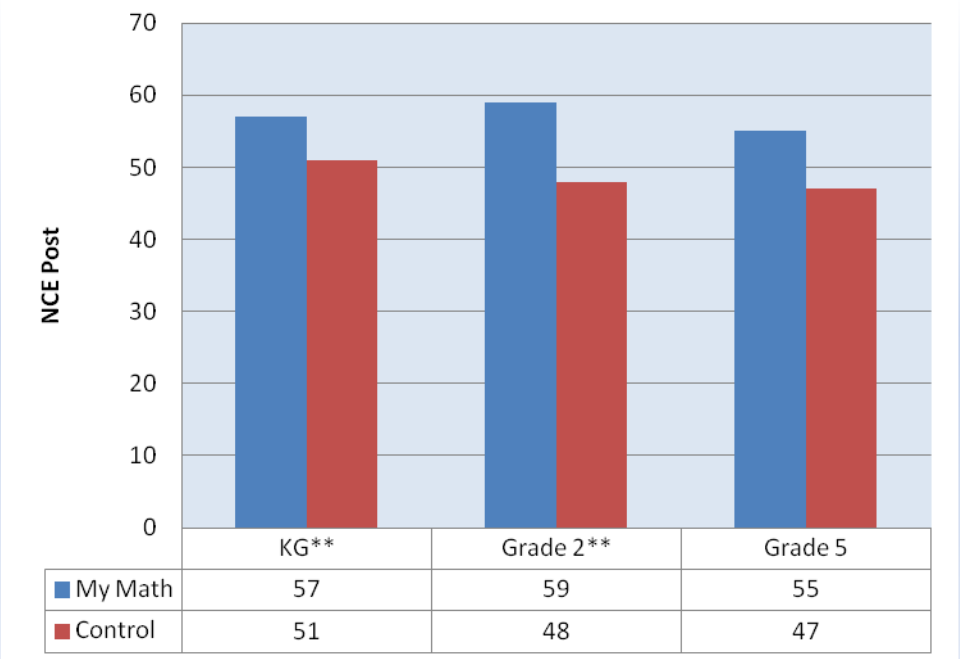
pretest scaled score, program condition was *not* found to be a statistically significant predictor of posttest scaled score (unstandardized regression coefficient $b = 3.72$, $t_{239} = .784$, $p = .434$). After controlling for students' pretest NCE score, the program condition was *not* found to be a statistically significant predictor of posttest NCE scores, (unstandardized regression coefficient $b = 2.59$ $t_{239} = 1.42$, $p = .16$). These findings did not change when gender, ethnicity, free/reduced lunch status, ESL/ELL status, or special education status variables were included in the statistical model.

Summary of Student Achievement

Figures 1 and 2 provide a summary of findings for all grade levels. Figure 1 compares the average performance of *McGraw-Hill My Math* and control students on the NCE posttest. The difference in scores for Kindergarten and second grade students is statistically significant, as indicated by the asterisks. The difference in scores for the fifth grade students was not found to be statistically significant.

Figure 1

Comparison of My Math and Control Students, NCE posttest

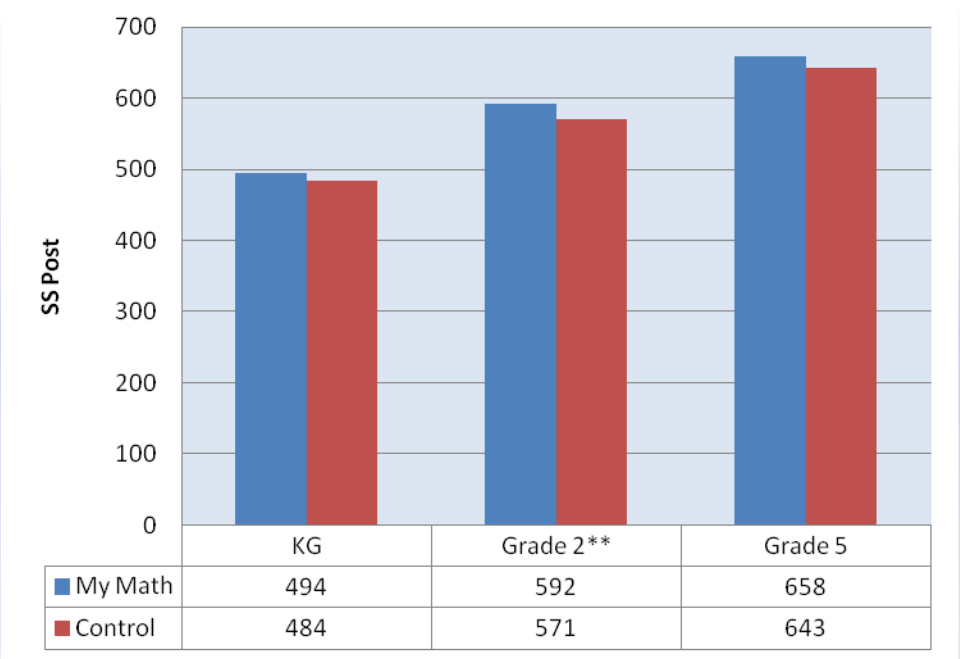


**Results statistically significant, $p < .05$

Figure 2 compares the average performance of *McGraw-Hill My Math* and control students on the NCE posttest. The difference in scores for second grade students is statistically significant, as indicated by the asterisks. The difference in scores for the Kindergarten and fifth grade students was not found to be statistically significant.

Figure 2

Comparison of My Math and Control Students, SS posttest



	KG	Grade 2**	Grade 5
My Math	494	592	658
Control	484	571	643

**Results statistically significant, $p < .05$

Teacher Feedback

Survey Responses

In May and after having several months of exposure to the program, teachers were asked to complete an online survey regarding their experiences and perceptions about *McGraw-Hill My Math*. Teachers were informed that the survey was administered and responses received by an independent researcher not directly employed by McGraw-Hill. Teachers also were aware that responses, whether positive or negative, had no bearing on their compensation provided at the end of the study.

The survey contained 49 items, 39 of which were constructed using a 5 point, Likert Scale with response selections ranging from *strongly agree* to *strongly disagree*. The remaining 10 items were constructed in an open-response format. While several questions were designed to solicit teachers' general perceptions of the program, others were more pointed and reflected specific program components. Open-ended responses were analyzed to determine the presence of consistent themes, across respondents. Ninety-five percent, or 20 of 21 teachers, completed and returned the survey within the requested time period.

Survey items and corresponding responses are grouped, by theme.

General Impressions

Teachers generally felt that the program is of high quality, with 18 of 20 respondents selecting *strongly agree* or *agree*. An overwhelming majority of respondents felt that preparation and instruction was manageable, with 95% (n=19) of respondents selecting *strongly agree* or *agree* for this statement. Initial impressions of the *Teacher Edition* were generally favorable, with about 70% (n=14) of respondents indicating that first impressions were positive; however, about 25% (n=5) of teachers reported feeling *neutral*, perhaps a reflection of lack of training. The order of presentation of topics was appropriate, according to 75% (n=15) of respondents. About 80% felt that *McGraw-Hill My Math* met their expectations, and 75% would recommend *McGraw-Hill My Math* to a colleague.

Program Relevance

Several items were constructed to determine *McGraw-Hill My Math's* perceived relevance to current academic content standards. About 90% of teachers (n = 18) selected either *strongly agree* or *agree* that *McGraw-Hill My Math* adequately reflects the *Common Core State Standards in Mathematics*, while one teacher each selected *neutral* and *disagree*. About 85% of teachers responded that *McGraw-Hill My Math* fully supports 21st Century Skills. There was general consensus that the examples provided in *My Math* are relevant to students' lives, with all teachers expressing agreement with this statement. About 85% of respondents felt that the *Standards for Mathematical Practice* are adequately covered. The majority of teachers felt that the lesson content is grade appropriate, with 90% of teachers selecting either *strongly agree* or *agree*.

Perceived Success

Generally, teachers and students were perceived to feel successful when using *McGraw-Hill My Math*; 85% (n=17) of teachers responded that they felt successful teaching the program, and 95% (n=19) of teachers responded that their students felt successful using the program. About 95% of teachers report that their students are fully engaged during instruction; similarly, 95% of teachers report that their students were able to independently complete learning activities.

Differentiated Instruction

When asked whether *McGraw-Hill My Math* meets the needs of diverse learners, 75% (n = 15) of teachers selected *strongly agree* or *agree*, while 15% (n = 3) selected *neutral*, and 10% (n = 2) selected *disagree*. About 40% of teachers felt that the program benefits those for whom English is a second language, and 60% reported feeling neutral; however, teachers selecting *neutral* indicated that they did not have a considerable ESL/ELL population at their school. About 55% (n=11) of teachers selected *strongly agree* or *agree*, indicating that they feel the program benefits special needs students, 40% (n=8) selected *neutral* (teachers indicated they did not have special needs students in their classrooms) and 5% (n=1) of teachers selected *disagree*. While many teachers felt they could not comment on the appropriateness of *McGraw-Hill My Math* for select subgroups of students, 80% (n=16) of teachers feel that *My Math* allows for scaffolding and differentiation of instruction to meet the diverse needs of learners.

Program Components

Several survey items were constructed to determine perceptions about individual program components. Fifteen of twenty (75%) teachers felt that the *My Math Words* page is an effective tool for determining student acquisition of vocabulary. Fewer teachers (n=13) selected *strongly agree* or *agree* to indicate that they felt the *Vocabulary Cards* form a critical piece of the program, and four teachers reported feeling 'neutral'. One-fourth of teachers (n=5) report that their students "always" use the corresponding *Vocabulary Cards*. Teachers also expressed mixed perceptions about the utility of the *Foldables* activity: eleven teachers (55%) report that they found the *Foldables* useful, five teachers (25%) report feeling 'neutral' and four teachers (20%) report that the *Foldables* were not particularly useful. While most teachers (n=16) felt that the *Student Interactive Text* provided ample room for students to record their responses, about 20% (n=4) of respondents felt that it did not. Twelve, or 60% of teachers indicate that the *Student Edition* is better than any other student edition they've used, and six (30%) teachers reported feeling 'neutral'

Technology

As stated previously, teachers did not immediately receive login codes for accessing digital resources; rather, technological features of the program became available at various times throughout the study and certainly well after teachers were entrenched in their teaching. While it was expected that responses regarding the technological features would perhaps be mixed, it was deemed important to solicit information about this component. About 55% of teachers found the digital platform *ConnectEd* easy to navigate, 35% felt 'neutral' (as they did not consistently access online features), and 10% experienced difficulty. About 30% report accessing online materials on a daily basis, and 60% responded that they did not. Upon follow-up, this was primarily a consequence of inconsistent access to computers. Several teachers reported that there is "only one computer in the classroom". About 35% of teachers selected *strongly agree* or *agree* that the online assessments were considered a valuable tool for assessing student knowledge, and 60% reported feeling neutral. Only 15% of teachers reported frequently using the online lesson planning feature, and 55% of teachers reported that they never use it.

Open-ended responses

When asked to provide four words to describe the program, about 72% of responses contained terms that indicated teachers/students found the program engaging or colorful (e.g., colorful, creative, engaging, fun). About 55% of responses included terms reflecting the ease in which the program can be implemented (e.g., easy to use, manageable, kid-friendly, teacher-friendly). About 40% of responses included terms reflecting the standards-based nature of the program (e.g., Common Core, Standards-Based).

When asked what teachers had hoped to find that they did not, the majority of teachers responded, "nothing". Kindergarten teachers frequently reported that they wanted to see additional manipulatives included as supplemental kits. Teachers were additionally asked to list the first and last resource or component teachers would be willing to relinquish. About 24% of teachers replied, 'nothing'. Other responses varied considerably. Resources teachers would be willing to give up include the *Vocabulary Cards* (18%), *Foldables* (11%), the online component (11%), and the *Teachers Edition* (5%). While many teachers felt that 'all' components are critical to the program, several teachers listed the online component and features (24%), the *Student Interactive Text* (18%), and the *Teachers Edition* (5%) as indispensible. When teachers were asked what they perceive are students' favorite parts of the program, the top three components include the *Student Interactive Text* (28%), the online games (28%), and the *Foldables/Vocabulary Cards* (28%). Other components include the problem solving activities (5%), the *Explore and Explain* feature (5%) and the *Practice Pages* (5%).

Teachers also provided their opinion regarding what they felt the *McGraw-Hill My Math* program does better than other programs they've used. When asked what *McGraw-Hill My Math* does better than others, 41% of teachers responded, "lots of practice"; 24% of teacher cited

"problem-solving"; 18% of teachers responded, "manipulatives and the tear-out pages", and 12% of teachers stated, "it's easy". Additional comments are provided, below:

> There are several things that this edition does better: **1.** I like the pre-test to see where the children are with the understanding of the concept before it is taught. **2.** I like the vocabulary cards that are provided. They are a good tool for the children to have and take home to use for practice. **3.** I like the way that the lessons are set up; providing for modeling, practice, and independent implementation of the skill.

> Many schools do not have access to a lot of manipulatives. This book has visuals/manipulatives in the text so that the students can use those to solve problems in hands-on ways.

> The student edition has real life application problems, has guided practice built right in so that the teacher doesn't have to make up more practice problems for the white board and presents the students with problem solving skills that are used repeatedly.

About half of the teachers reported that the differentiated instruction component may need modification. As one fifth-grade teacher reported:

> I feel that the differentiation component (specifically the assessments) could be better. The enrichment activities do not require higher level thinking and the *reteach* activities do not meet the needs of students who are "approaching." They are great for additional practice, but not for a student who truly "does not get it."

The resources or activities most frequently used to differentiate instruction include the *Am I Ready?* Pre-assessment (35% of respondents), followed by the online worksheets/games (18% of respondents). Other activities included extra homework practice, *Reteach/On My Own* pages, ideas obtained from the *Teachers Edition*, and practice using the *Foldables/Vocabulary Cards*.

For ELL students, most teachers found the *Vocabulary Cards* helpful:

> I really liked the vocabulary cards. They are helpful not only to the students, but are helpful to the parents. When I send home the Homework pages, the parents can refer to the vocabulary cards, as well. I also like the pictures/photographs that are used in the student edition. It provides clarity for these students.

Finally, teachers were provided opportunity to provide final comments describing their experiences with *McGraw-Hill My Math* over the course of the school year:

> This program provides students with exciting and interactive ways to learn math. Students are provided with multiple opportunities to practice and ultimately master a concept. There are vocabulary cards and a bilingual glossary included with this program.

The assessments that were provided are leveled according to the ability of students. It was easy to implement and students enjoyed using this program.

It is hard to manage with kindergarten students since the program in divided into 2 huge workbooks that are hard to tear out. Each lesson does not give enough problems to adequately practice a new skill. The lessons appear very redundant.

Personally, I liked using the program. I liked the way that it was scaffolded and taught things in a nice progression. I do feel that not all students can understand number sense to 20 at this age, but having them at least handle that many manipulatives gives them an idea of what that is. I also liked the way that the lessons were labeled with the Core Curriculum standards. It is helpful to see how the lesson aligns with what needs to be taught.

I have described it as a program I do not want to live without next year. The layout and compatibility for Smartboard is my favorite. Lesson planning is easier than ever!

McGraw-Hill My Math is a very effective math program. It is teacher friendly and the students are actively engaged while using it. It has a lot of nice online support and extended activities and games. Once the program has been re-edited (corrected) for publication, it will be the best math program I have seen available for second graders.

This program is an in depth approach to teaching the fundamentals of math in real life situations. Students are given ample opportunity to problem solve, reflect on their approach to working out problems and helps students realize that there are many different ways to solve the same problem.

Slower moving than I am used to, but students learn content more deeply than in the past.

Interviews

Eleven teachers were randomly selected to participate in an interview, conducted either on-site during the classroom observation or off-site via telephone. Ten teachers completed the interview; one teacher was not available during an extended period over the summer, and could not participate.

The three Kindergarten teachers interviewed expressed similar reactions to the program. Teachers found the program very visually appealing, and reported that it offers "a whole lot of practice" and is "very strong in number sense". Teachers were generally impressed with "what kids were getting out of the program". One Kindergarten teacher stated that the program was "very easy to use, even for a first-year teacher". While all Kindergarten teachers reported that they very much liked the program and found it relevant, there were some recommendations for improvement. All teachers often experienced difficulty when Kindergarten students were asked to remove the perforated pages in the *Student Interactive Text*. One teacher circumvented this difficulty by removing needed pages and stapling them together the day before each lesson, but

she reported that “this was very time consuming and it seems like it shouldn’t be necessary”. All teachers expressed an interest in having *blackline masters* be made available (they were not available at that time) and a set of manipulatives be included with the program. All teachers emphatically stated that technology training is vital.

Comments made by second and fifth grade teachers generally echoed that of Kindergarten teachers. Nearly all teachers expressed that they feel the program is very easy to use and implement, and is highly engaging for students. Many students had never “had their own textbook to use”, and as such found the *Student Interactive Text* highly motivating. All teachers responded that they noticed “an increase in problem-solving ability” and that “students use to hate word problems, but they don’t complain as much about them now’. The recommendation most commonly reported by second and fifth grade teachers concerned the *Enrichment* pages, which “do not really require higher-level thinking skills”. One fifth grade teacher reported that she felt the *Enrichment* pages were merely “skill and drill”, and lacked more drawn-out story problems that require higher-level thinking and skill application. The difficulty students experienced tearing out pages was expressed by some, but not all, second grade and fifth grade teachers. Second and especially fifth grade teachers felt that more room was needed for students to record their responses, as “fifth grade writing is very large and loopy”. Other recommendations include suggestions for pacing (some teachers felt the pacing was too slow) and that technology training would have been very useful and helpful.

Discussion

The purpose of this evaluation was to determine the effects of *McGraw-Hill My Math* on student achievement, as well as on the attitudes and perceptions of teachers using the program. Thus, the study serves as a summative as well as formative evaluation.

There were several factors in this study that no doubt impact the interpretation of the *Terra Nova* data, and any conclusions regarding the program effects should be considered within this context. As stated, teachers received all print materials in mid-October, with implementation occurring in late October. As the program was quite new and the technological features were only just becoming available, many teachers did not receive their digital access until November. Since it was desired to determine the program's effect as implemented in a natural setting and since formal training components were not available, teachers did not receive training in implementing either the print or technological components. This is acceptable in terms of formative evaluation, (as part of the reason for this evaluation is to determine training needs), but it is a factor that impacts the quantitative findings. Many teachers openly admitted that they either did not access online features at all, or experienced frustration accessing/using features and "gave up" trying. As technology is a critical component of *McGraw-Hill My Math*, this needs to be taken into consideration when interpreting the findings.

While the difference in pre-posttest scores did not achieve statistical significance for Kindergarten and fifth-grade, findings were statistically significant for second-grade, and this is of particular interest. While most teachers teaching the program were the only teachers within their respective schools to use the program, it should be noted that five of the second-grade teachers (three in the *McGraw-Hill My Math* group, two in the control group) were housed within the same school. A visit to the site and interviews with these teachers revealed that there existed a great deal of support and camaraderie among them. When faced with challenges or questions regarding implementing certain features of the program, teachers provided support to each other and suggested ideas for circumventing such challenges (e.g., tearing out workbook pages in advance and stapling them together for Kindergarten students who may have experienced difficulty tearing pages out on their own). Further, this group was particularly careful to implement the program with fidelity (e.g., they consistently utilized the *Vocabulary Cards* and *Foldables*, whereas a few teachers did not regularly do so). After controlling for initial performance on the pretest, second grade students using *McGraw-Hill My Math* outperformed the control group by about 14 scaled score units and 6 NCE units, on average. It seems logical to conclude that fidelity of implementation and/or collaboration and support may have impacted these findings in a positive way. Another factor that may impact conclusions rests with the alignment of *McGraw-Hill My Math* which by design addresses the *Common Core Standards*, to the assessment instrument (e.g., *Terra Nova*) which is not currently aligned to the *Common Core Standards*. Should this study be replicated, it is recommended that training be

included prior to program implementation, and the assessment instrument be closely aligned with the skills covered in the program.

Teachers' perceptions of the program were largely positive. A majority felt that *McGraw-Hill My Math* adequately addresses the *Common Core Standards for Mathematics* and supports 21st Century skills. Overwhelmingly, teachers felt that *McGraw-Hill My Math* was very easy to implement – even for first-year or inexperienced teachers – but many did not express the same degree of confidence when accessing technology. Nearly all teachers indicated that training on the technology component would be very helpful. In regard to student engagement, 95% of teachers reported that their students were fully engaged during instruction. Teachers report that students liked, in particular, the *Student Interactive Text*, as being able to "write in a math book" was considered by some students as "very cool".

Teachers varied, to some degree, in their dislikes; for example, some teachers found the *Vocabulary Cards* not particularly useful, while others felt they were indispensible. The recommendations that follow represent the most commonly cited suggested modifications reported by teachers at all grade levels.

Recommendations for Program Improvement[1]

Based on responses provided during the interviews and surveys, the following are recommendations for program improvement:

- Addition of Blackline masters
- Modification of the *Enrichment* pages to reflect higher-level thinking skills
- Training to address technology – a must
- Suggestions for pacing
- Examples and steps on the *Re-teaching* pages
- Examine vocabulary to ensure it is not confusing for students; for example, on the fifth grade tutorial, "now we are going to *annex* the zero"
- Add more room for students to write (fifth graders' writing is large and loopy)
- Add more basic computation skills to front page
- Put *Vocabulary Cards* on cardstock; paper is too flimsy
- Reconsider the *Foldables* for Kindergarten – too difficult
- Add pages for fluency practice
- Provide suggestions for how to use the *Graphic Novels*
- Include a statement of objective or skill at the bottom of the page
- Use different colors for Vol. 1 and 2 to help differentiate
- For Kindergarten, include shape identification activities at the beginning of the program
- For Kindergarten, include patterning activities

[1] Many of these suggestions have been addressed as the program was further developed

References

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., and Brenwald, S. (2008). Highlights from TIMMS, 2007: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context. (NCES 2009-001 Revised). *National Center for Education Statistics, Institute of Education Sciences, U.S., Department of Education*. Washington, D.C.

National Center for Education Statistics (2011). The nation's report card: Mathematics 2011. (NCES 2012-458). *Institute of Education Sciences, U.S. Department of Education,* Washington, D.C.

National Mathematics Advisory Panel (2008). Foundations for success: The final report of the national mathematics advisory panel. U.S. Department of Education: Washington, DC.