Examining the Impact of Imagine Math® on MCAP Math Scores: A Quasi-Experimental Study

Abstract

This study examines the impact of Imagine Math on student performance on the Maryland Comprehensive Assessment Program (MCAP) using a quasi-experimental design with propensity score matching (PSM). While the most rigorous analytical approach (Research Question 1; RQ1) found no significant positive effects, further analysis revealed that low student engagement likely influenced this result — 25% of students in the treatment group completed zero lessons, and average lesson completion was below recommended thresholds.

When restricting the analysis to students who used Imagine Math to a greater extent (RQ2, RQ3), significant positive associations emerged, indicating that students who completed the recommended number of lessons in Imagine Math performed better on the MCAP. Additionally, analysis of usage bands (RQ4) showed a dose-response relationship, with greater lesson completion linked to higher achievement gains.

These findings highlight the importance of program engagement in achieving academic growth. While direct comparison between program users and non-users showed no advantages for Imagine Math users, focusing on students who used the program with fidelity showed clear gains above program non-users. Ultimately, increased use of Imagine Math appears to improve performance on the MCAP and educators are encouraged to maximize student gains by implementing Imagine Math at the recommended dosages.

Introduction

As digital learning tools become increasingly prevalent in K–12 education, school districts are investing significant resources in programs designed to enhance student achievement. Mathematics programs like Imagine Math are particularly appealing because they provide adaptive, personalized learning experiences intended to supplement classroom instruction. Given the widespread adoption of such programs, it is critical for school districts to evaluate their impact and determine how to maximize their effectiveness for all students.

Understanding the relationship between program usage and student outcomes is essential for making informed decisions about instructional technology. School leaders need clear, data-driven insights to identify the conditions under which they are most effective.

This study examines the impact of Imagine Math on student performance on the Maryland Comprehensive Assessment Program (MCAP) in an urban public school district by

addressing five research questions (see below). By analyzing student engagement levels and performance outcomes, we aim to provide actionable insights that help educators optimize program implementation and improve student success.

Research Questions

RQ1. Do students who use Imagine Math demonstrate better performance on the MCAP compared the students who do not use Imagine Math?

RQ2. Do students use who use Imagine Math with fidelity demonstrate better performance on the MCAP compared to students who do not use Imagine Math?

RQ3. Do students who use Imagine Math with at least 80% fidelity demonstrate better performance on the MCAP compared to students who do not use Imagine Math?

RQ4. What is the MCAP performance of students grouped by Imagine Math usage bands according to time in the program and lessons passed?

RQ5. How do associations between Imagine Math and student achievement vary by student demographics (English learner classification, race/ethnicity, students with disabilities, and prior achievement)?

Methods

Study Design and Participants

A quasi-experimental design was used, leveraging propensity score matching (PSM) to equate students in the treatment and control groups. The study included students from Grades 6 through 8 in an urban public school district, with some using Imagine Math and others not. To address the first research question, the treatment group was initially comprised of all students who had access to and actively used Imagine Math, while the control group consisted of students who did not utilize the intervention. The second and third research questions were addressed by limiting the treatment sample to students who fit the appropriate usage profiles (RQ2: passing 30 lessons or more, RQ3: passing 24 lessons or more) and conducting PSM for each analysis to ensure a comparable control group of students who did not use the program at all (see next section).

Propensity Score Matching

To minimize selection bias, students were statistically matched based on their previous year's math scores, grade level, gender, race, English language (EL) classification, socioeconomic status (FARMS), disability status (SWD), and home language. Post-match balance checks were conducted to ensure that the resulting matched groups were statistically comparable on these variables.

Results

Matching Analysis

Balance checks indicated successful matching across all demographic and academic covariates for Research Questions 1–3, with standardized mean differences (SMD) below the threshold of 0.25 for most variables, indicating that the study groups were comparable for each analysis (see Appendix A). Multivariate regression analyses were then utilized to control for any remaining differences between the study groups and compare academic performance.

RQ1. Do students who use Imagine Math demonstrate better performance on the MCAP compared the students who do not use Imagine Math?

A regression analysis was conducted to compare the Spring 2023 MCAP scores of treatment and control groups while controlling for Spring 2022 scores, grade level, gender, race/ethnicity, EL classification, socioeconomic status, disability status, and home language. Results revealed that the use of Imagine Math was not associated with improved MCAP Math scores. Importantly, these initial findings need to be interpreted within the broader context of student engagement and fidelity of program implementation. Specifically, the average use of Imagine Math for the defined treatment group was 6.22 hours (*SD* = 8.16) and 5.94 lessons passed (*SD* = 13.64). This falls far below Imagine Learning's recommended levels of 30 hours and 30 lessons passed. Other significant predictors of MCAP Math scores included prior-year math scores, ELL status, and socioeconomic status. (See Appendix B for full results of the regression analysis.)

To assess the practical significance of these results, difference scores (posttest scores – pretest scores) were compared between groups. Difference scores differed by less than a tenth of a point between treatment and control groups (*M* = 1.34, *SD* = 14.50 vs. *M* = 1.25, *SD* = 13.88, respectively).

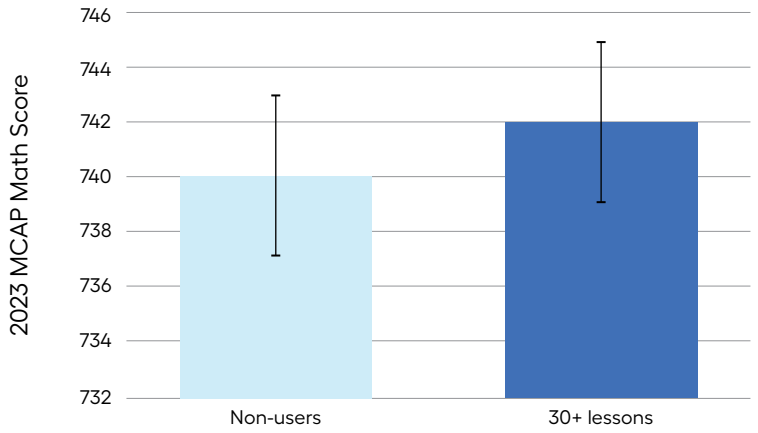
RQ2. Do students use who use Imagine Math with fidelity demonstrate better performance on the MCAP compared to students who do not use Imagine Math?

Low average use of the Imagine Math program may have reduced the sensitivity of the analyses in RQ1 to sufficiently measure the impact of the program on student MCAP performance. As such, for RQ2, a similar regression analysis was conducted to compare the Spring 2023 MCAP scores of treatment and control groups while controlling for Spring 2022 scores, grade level, gender, race/ethnicity, EL classification, socioeconomic status, disability status, and home language. However, this analysis examined the relationship between Imagine Math usage with fidelity (i.e., passing at least 30 lessons) and MCAP Math scores. Results revealed a statistically significant positive association ($\beta = 2.41$, $p < 0.01$). This indicates that

students who met the recommended usage threshold demonstrated higher MCAP Math scores compared to students who did not use the program. (See Appendix B for full regression results). See Figure 1 for the adjusted average 2023 MCAP Math scores of each group.

Without controlling for covariates, the fidelity treatment group gained more than two and a half points more than the control group, (M = 1.88, SD = 11.32; M = -0.73, SD = 11.32, respectively). This difference was significant, $t(726) = 3.07$, $p < .01$.

Figure 1. Adjusted average 2023 MCAP scores of students who passed 30 or more lessons in Imagine Math compared to students who did not use the program during the 2022–2023 school year.



Note: Means are adjusted for the following covariates: Spring 2022 scores, grade level, gender, race/ethnicity, EL classification, socioeconomic status, disability status, and home language. Adjusted means are significantly different from each other ($p < .01$). Error bars represent standard error.

RQ3. Do students who use Imagine Math with at least 80% fidelity demonstrate better performance on the MCAP compared to students who do not use Imagine Math?

Results from RQ2 indicate that using Imagine Math with fidelity leads to improved student outcomes. Further analyses were conducted to explore the bounds of this relationship. Additional regression analysis assessing the relationship between near-fidelity (at least 80% fidelity) Imagine Math usage (i.e., passing at least 24 lessons) and MCAP Math scores, while controlling for multiple covariates, indicates a significant positive association ($\beta = 1.88$, $p < 0.01$). Students who used the program with at least 80% fidelity performed significantly better on the MCAP Math assessment compared to students who did not use the program. These results reinforce the importance of sustained engagement with Imagine Math. (See Appendix B for full results of the regression analysis.)

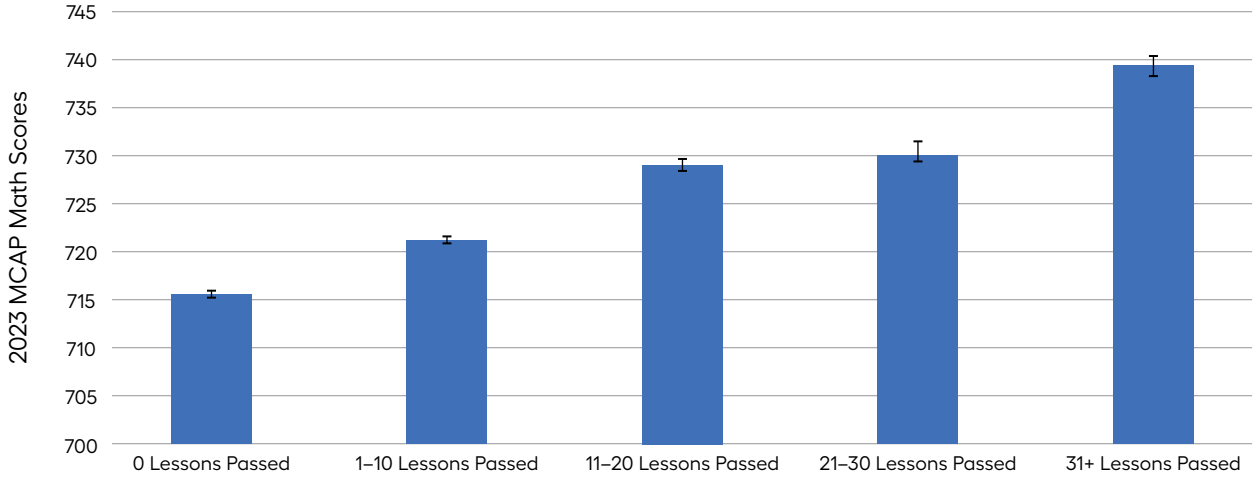
The 80% fidelity treatment group gained more than two points more than the control group, (M = 1.95, SD = 11.44; M = -0.18, SD = 11.74, respectively). This difference was significant, $t(994) = 2.90$, $p < .01$.

RQ4. What is the MCAP performance of students grouped by Imagine Math usage bands according to time in the program and lessons passed?

To directly explore the association between increased use of Imagine Math and MCAP Math performance, additional analyses examined MCAP Math performance across different Imagine Math usage bands based on lessons passed. Specifically, all Imagine Math users were organized into the following groups based on how many lessons they passed in Imagine Math: 0 lessons passed, 1–10 lessons passed, 11–20 lessons passed, 21–30 lessons passed, and 31 or more lessons passed. ANCOVA was used to compare the Spring 2023 scores of each group, controlling for the following variables: Spring 2022 scores, grade level, gender, race/ethnicity, EL classification, socioeconomic status, disability status, and home language.

The ANCOVA results indicated a statistically significant difference in performance across groups (F (4, 8797) = 321.87, p < .01, η^2 = .13), suggesting that students who engaged with Imagine Math at different levels demonstrated varying levels of improvement in their MCAP scores. The adjusted mean scores show that students who did not pass any lessons had the lowest Spring 2023 MCAP Math scores (M = 715.01, SD = 14.25), while students who passed 31 or more lessons had the highest mean difference score (M = 744.10, SD = 16.99). See Figure 2 for a graph comparing mean difference scores. Post-hoc Bonferroni comparisons revealed that all groups were significantly different from one another (p < .01), with the exception of the 11–20 and 21–30 lessons passed groups (p = .71). Overall, these findings indicate that higher engagement with Imagine Math, particularly among students passing 31 or more lessons, is associated with greater MCAP performance.

Figure 2. Adjusted average 2023 MCAP Math scores of Imagine Math students grouped by the number of lessons passed in Imagine Math.



Note: Means are adjusted for the following covariates: Spring 2022 scores, grade level, gender, race/ethnicity, EL classification, socioeconomic status, disability status, and home language. All groups are statistically different from one another (ps < .01) except for the 11–20 and 21–30 lessons passed groups (p = .71). Error bars represent standard error.

RQ5. How do associations between Imagine Math and student achievement vary by student demographics (English learner classification, race/ethnicity, students with disabilities, and prior achievement)?

The relationship between lessons passed in Imagine Math and student achievement was analyzed separately for each category of each demographic variable: multi-language learner classification, race/ethnicity, special education classification, and prior achievement. Only students who logged time in Imagine Math were included in these analyses. Because the distribution of lessons passed was positively skewed, a logarithmic transformation was applied to normalize the data before conducting regression analyses. This transformation ensures that extreme values do not disproportionately influence the results while allowing for a more interpretable relationship between Imagine Math engagement and student achievement.

English learner (EL) classification. The results indicate a positive association between lessons passed in Imagine Math and MCAP Math scores across all EL classifications. EL students demonstrated a statistically significant improvement in MCAP scores ($\beta = .22$, $p < .01$), suggesting that EL students who used Imagine Math experienced notable gains. Students classified as non-EL showed a similar positive effect ($\beta = .24$, $p < .01$), as well as students who have exited EL services ($\beta = .26$, $p < .01$), suggesting that Imagine Math usage had a strong association with improved performance for all groups.

Race/ethnicity. The groups with a sufficient number of students to conduct regression analyses were Black or African American ($n = 6521$), Hispanic/Latino ($n = 1571$), and White ($n = 537$). The results indicate a positive association between Imagine Math usage and MCAP Math scores across all three groups. Black or African American students showed a significant and positive effect ($\beta = .23$, $p < .01$), similar to that observed by Hispanic/Latino students ($\beta = .24$, $p < .01$). White students experienced the weakest association ($\beta = .16$, $p < .01$), but still positive and significant.

Students with disabilities. Categories of students include those with a 504 plan ($n = 332$), an IEP ($n = 1415$), and no disability ($n = 7063$). All groups demonstrated a similar positive association between Imagine Math usage and MCAP Math scores (504: $\beta = .26$, $p < .01$; IEP: $\beta = .20$, $p < .01$; No disability: $\beta = .24$, $p < .01$).

Prior achievement. Four categories of students include: Beginning Learner, Developing Learner, Proficient Learner, and Distinguished Learner. Only three groups, Beginning Learners ($n = 5510$), Developing Learners ($n = 2833$), and Proficient Learners ($n = 449$) had enough students to report on results. Developing Learners saw the strongest association between Imagine Math usage and MCAP Math scores ($\beta = .22$, $p < .01$), followed by Proficient Learners ($\beta = .17$, $p < .01$) and Beginning Learners ($\beta = .10$, $p < .01$). These results suggest that Imagine Math had an association with improved performance for all learner groups.

Discussion

The findings of this study highlight the complex but clear relationship between Imagine Math usage and student achievement on the MCAP. While comparing all Imagine Math users to all Imagine Math non-users may follow the common industry practice of measuring effects among the intent-to-treat sample, this study demonstrates that such an approach may not always be the most appropriate given the observed patterns of program engagement in the study sample. During the initial analysis, it was found that although all students who logged into Imagine Math were included in the treatment group, 25% of those students completed zero lessons. Furthermore, on average, students completed far fewer lessons than recommended. This extremely low level of engagement among a substantial portion of the treatment group likely diluted any potential effects of the program when examining its overall impact (RQ1).

The results for RQ1 indicated that students in the Imagine Math treatment group performed slightly worse on the MCAP than their matched peers who did not use the program. However, this finding must be interpreted with caution. The negative effect observed is unlikely due to the program itself but rather to the inclusion of students who had access to Imagine Math but did not meaningfully engage with it. This is evident when examining the difference scores of the two groups: the practical differences in growth between groups was negligible and insignificant. This issue underscores the importance of considering fidelity of implementation when evaluating educational interventions.

In contrast, the analyses for RQ2 and RQ3, which focused on students who used the program to greater degrees, yielded positive and statistically significant associations between Imagine Math usage and MCAP performance. These findings suggest that when students engage with Imagine Math as intended — meeting the recommended lesson completion thresholds — they experience measurable benefits in their math achievement.

Further supporting this interpretation, the results of RQ4 demonstrated a clear dose-response relationship, with students in higher Imagine Math usage bands showing greater improvements in their MCAP scores. The ANCOVA analysis revealed significant differences in performance between students who completed fewer lessons and those who completed more lessons, reinforcing the conclusion that program impact is contingent on sufficient engagement.

Taken together, these findings emphasize the need for careful consideration of student engagement when evaluating the efficacy of digital learning tools. While the use of a PSM approach in RQ1 was methodologically sound, it may not have provided the most meaningful

estimate of Imagine Math's impact given the large proportion of students with minimal or no practical usage. Efforts should be made to understand the barriers to student participation in Imagine Math and to identify strategies for increasing fidelity of implementation.

Subgroup analyses further offer important insights into how Imagine Math may differentially support student learning based on key demographic and academic characteristics. The positive associations observed across all English learner classifications suggest that the program can be a valuable resource for supporting language-diverse learners. Similarly, the program was positively associated with improved outcomes across the three largest racial/ethnic groups, with the strongest effects observed for Black or African American and Hispanic/Latino students. Students with and without disabilities all demonstrated comparable gains, indicating that Imagine Math may be accessible and beneficial regardless of disability classification. Additionally, the program was associated with improved outcomes across levels of prior achievement. Taken together, these findings suggest that Imagine Math is broadly applicable across student populations.

In conclusion, this study provides evidence that Imagine Math is positively associated with student achievement in mathematics, particularly when students engage with the program at recommended levels. Using a quasi-experimental design and propensity score matching methods, the analysis found that students who used Imagine Math with fidelity demonstrated significantly greater gains on the MCAP Math assessment compared to their peers. These effects were consistent across multiple demographic groups. While overall usage of the program was lower than recommended, the results suggest that Imagine Math has the potential to support equitable math achievement when implemented with fidelity. Future work should focus on strategies to increase meaningful engagement with the program and examine how Imagine Math can be integrated into instruction to maximize its benefits for all students.

Appendix A

Table A1. Baseline equivalence for all users vs. all nonusers in final matched sample for RQ1.

Group	Subgroup	Comparison Students	Imagine Math Students	*p*-value	Standardized Mean Difference (SMD)
n		2808	2808		
Average (SD) Spring 2022 MCAP Math Score		722.66 (16.62)	718.92 (17.53)	<.001	0.219
Grade Level	6	1032	1032	>.999	<0.001
	7	987	987		
	8	789	789		
Gender	Female	1495	1442	.165	0.038
	Male	1313	1366		
Students with Disabilities	504	150	139	<.001	0.234
	IEP	462	729		
	No disability	2196	1940		
English Language Learner	Exited	92	118	.168	0.050
	No	2596	2564		
	Yes	120	126		
Economically Disadvantaged	No	959	959	.251	0.031
	Yes	1849	1849		
Race/Ethnicity	American Indian/ Alaskan Native	5	9	.008	0.111
	Asian	8	13		
	Black or African American	2298	2180		
	Hispanic/Latino	277	324		
	Multiracial	31	34		
	Native Hawaiian/ Pacific Islander	2	2		
	White	187	246		
Home Language	Arabic	9	11	.249	0.070
	Chinese	2	4		
	English	2565	2523		
	French	6	10		
	Spanish	223	259		
	Swahili	1	1		
	Wolof	1	0		
	Yoruba	1	0		

Table A2. Baseline equivalence for fidelity users (30+ lessons passed) vs. nonusers in final matched sample for RQ2.

Group	Subgroup	Comparison Students	Imagine Math Students	*p*-value	Standardized Mean Difference (SMD)
n		364	2364		
Average (SD) Spring 2022 MCAP Math Score		739.04 (15.03)	739.36 (15.13)	.776	0.021
Grade Level	6	202	202	>.999	<0.001
	7	141	141		
	8	21	21		
Gender	Female	194	185	.553	0.050
	Male	170	179		
Students with Disabilities	504	10	15	.366	0.105
	IEP	12	17		
	No disability	342	332		
English Language Learner	Exited	31	33	.881	0.037
	No	322	318		
	Yes	11	13		
Economically Disadvantaged	No	187	195	.603	0.044
	Yes	177	169		
Race/Ethnicity	American Indian/ Alaskan Native	1	1	.872	0.100
	Asian	3	5		
	Black or African American	236	221		
	Hispanic/Latino	54	57		
	Multiracial	9	9		
	White	61	71		
Home Language	Arabic	0	2	.390	0.151
	Chinese	1	3		
	English	317	310		
	Spanish	46	48		
	Swahili	0	1		

Table A3. Baseline equivalence for 80% fidelity users (24+ lessons passed) vs. nonusers in final matched sample for RQ3.

Group	Subgroup	Comparison Students	Imagine Math Students	*p*-value	Standardized Mean Difference (SMD)
n		498	498		
Average (SD) Spring 2022 MCAP Math Score		737.24 (15.69)	737.27 (15.38)	.974	0.021
Grade Level	6	270	270	>.999	<0.001
	7	192	192		
	8	36	36		
Gender	Female	253	254	>.999	0.004
	Male	245	244		
Students with Disabilities	504	14	22	.375	0.089
	IEP	27	29		
	No disability	457	447		
English Language Learner	Exited	42	41	.658	0.058
	No	443	439		
	Yes	13	18		
Economically Disadvantaged	No	248	256	.657	0.032
	Yes	250	242		
Race/Ethnicity	American Indian/ Alaskan Native	1	1	.607	0.121
	Asian	5	7		
	Black or African American	345	318		
	Hispanic/Latino	66	75		
	Multiracial	10	14		
	White	71	83		
Home Language	Arabic	3	2	.744	0.089
	Chinese	2	1		
	English	435	431		
	Spanish	57	64		
	Swahili	1	0		

Appendix B

Table B1. Regression Results for RQ1: Do students who use Imagine Math demonstrate better performance on the MCAP compared the students who do not use Imagine Math?

Coefficient	Estimate	Standard Error	*p*-value
Imagine Math User Indicator	-1.13	0.34	<.01
Intercept	284.36	9.61	<.01
Spring 2022 MCAP Score			<.01
Grade			
7	-1.44	0.40	<.01
8	3.80	0.43	<.01
English Language Learner Indicator			
N	-3.29	1.65	.05
Y	-6.94	1.20	<.01
Gender	-0.82	0.34	.02
Student with Disability Indicator			
IEP	-4.08	.82	<.01
No Disability	1.97	0.77	<.01
Economically Disadvantaged Indicator	2.69	0.37	<.01
Race/Ethnicity			
Asian	10.63	4.68	.02
Black or African American	-4.96	3.34	.14
Hispanic/Latino	-3.84	3.47	.27
Multiracial	-0.83	3.67	.82
Native Hawaiian/Pacific Islander	-11.34	7.16	.11
White	3.92	3.39	.25
Home Language			
Chinese	-7.02	6.49	.28
English	5.78	3.16	.07
French	8.38	4.26	.05
Spanish	7.45	3.10	.02
Swahili	21.42	9.32	.02
Wolof	19.69	12.87	.13
Yoruba	27.25	12.86	.03

Table B2. Regression Results for RQ2: Do students use who use Imagine Math with fidelity demonstrate better performance on the MCAP compared to students who do not use Imagine Math?

Coefficient	Estimate	Standard Error	*p*-value
Imagine Math User Indicator	2.41	0.82	<.01
Intercept	106.25	27.31	<.01
Spring 2022 MCAP Score	0.87	0.03	<.01
Grade			
7	3.86	0.87	<.01
8	8.84	0.82	<.01
English Language Learner Indicator			
N	-3.11	3.19	.333
Y	3.56	2.69	.19
Gender	0.40	0.83	.63
Student with Disability Indicator			
IEP	-0.35	3.07	.91
No Disability	-1.82	2.27	.42
Economically Disadvantaged Indicator	1.99	0.93	<.01
Race/Ethnicity			
Asian	-5.12	9.70	.60
Black or African American	-9.86	7.83	.21
Hispanic/Latino	-9.72	8.15	.23
Multiracial	-5.14	8.25	.53
White	-5.06	7.91	.52
Home Language			
Chinese	-10.52	11.21	.35
English	1.41	8.03	.86
Spanish	-0.80	8.25	.92
Swahili	14.32	13.64	.29

Table B3. Regression Results for RQ3: Do students who use Imagine Math with at least 80% fidelity demonstrate better performance on the MCAP compared to students who do not use Imagine Math?

Coefficient	Estimate	Standard Error	*p*-value
Imagine Math User Indicator	1.88	0.70	<.01
Intercept	106.23	22.17	<.01
Spring 2022 MCAP Score	0.86	0.03	<.01
Grade			
7	4.58	0.74	<.01
8	9.74	1.44	<.01
English Language Learner Indicator			
N	-2.48	2.74	.37
Y	-1.81	2.36	.44
Gender	-0.39	0.71	.58
Student with Disability Indicator			
IEP	-1.91	2.40	.43
No Disability	-0.71	1.90	.71
Economically Disadvantaged Indicator	1.94	0.79	.01
Race/Ethnicity			
Asian	-3.18	8.69	.71
Black or African American	-8.72	7.82	.26
Hispanic/Latino	-8.49	8.14	.30
Multiracial	-3.71	8.13	.65
White	-4.12	7.90	.60
Home Language			
Chinese	-5.84	8.91	.51
English	4.42	5.48	.42
Spanish	4.00	5.54	.47
Swahili	0.81	12.19	.94