DreamBox Math (2021-22)

Study Type: ESSA Evidence Level II

Prepared for:
DreamBox Learning

Prepared by LearnPlatform:
Meetal Shah, Ph.D., Senior Researcher, LearnPlatform
Alexandra Lee, Ph.D., Researcher, LearnPlatform

May 19, 2023

EXECUTIVE SUMMARY

DreamBox Learning contracted with LearnPlatform, a third-party edtech research company, to examine the impact of DreamBox Math usage on student math outcomes. LearnPlatform designed the study to satisfy Level II requirements (Moderate Evidence) according to the Every Student Succeeds Act (ESSA).

Study Sample, Measures, and Methods

This study occurred during the 2021-22 school year and included 13,589 K-8 students from across 39 schools in one district. Less than five percent of the K - 6 sample (*n* = 9,555) were true non-users[1] of DreamBox Math therefore, researchers split the analytic approach. Researchers used a correlative, treatment-only study for the K – 6 sample and a comparative study for the grades 7 – 8 sample (*n* = 4,034). In terms of demographics, the district was predominantly white (81%), and for the remainder of the students the racial breakdown was as follows: 2% Asian, 3% Black, 13% Hispanic, and 1% Other. Fifteen percent of the students had individualized educational programs (IEPs) and six percent had English language learner (ELL) status. The district did not provide data on students' socioeconomic status (e.g., low-income indicator, or free and reduced lunch status).

Researchers used DreamBox usage data and Houghton Mifflin Harcourt's (HMH) Math Inventory™ Quantile® scores to provide insights into DreamBox Math implementation and evidence about potential impacts of DreamBox Math on student learning outcomes.

Researchers used a variety of quantitative analytic approaches to answer the research questions. First, researchers used descriptive statistics to examine participant characteristics and support implementation analyses. Researchers then conducted linear regressions and partial correlations to examine how DreamBox Math use related to student math outcomes from beginning- to end-of-year. The analyses included student-level covariates to control for potential selection bias. Researchers also examined whether there were any differences between students who used DreamBox Math during the 2021–22 school year and students who did not use the program using linear regression analysis. In addition, researchers calculated standardized effect sizes (Hedge's *g*) to determine the magnitude of changes in student outcomes.

[1] Students who have had zero total use in terms of total lessons and time (in minutes) on the platform.

Student Outcomes

- Among K – 8 students who had ELL status, those who completed more lessons per week in DreamBox and those who spent more time (minutes per week) in DreamBox Math had higher spring Math Inventory™ scores.
- Grades 7 – 8 students who used DreamBox Math had statistically significant higher Quantile® scores on Math Inventory™ at the end of the year than students who did not use the program.
- Grades 7 – 8 students who completed at least 1.5 lessons per week in DreamBox Math had higher end-of-year Math Inventory™ scores compared to students who completed less than 1.5 lessons.
- Grades 7 – 8 students who spent at least 38 minutes per week in DreamBox Math had higher end-of-year Math Inventory™ scores compared to students who spent less than 38 minutes in the platform.
- For all grades from K – 6, students who completed more DreamBox Math lessons per week had higher spring Math Inventory™ scores. These were statistically significant relationships (*p values* < .05) for students in Grades 1 – 6.
- K – 6 students who completed at least 2.4 lessons per week in DreamBox Math had higher end-of-year Math Inventory™ scores compared to students who completed less than 2.4 lessons.
- K – 6 students who spent at least 58 minutes per week in DreamBox Math had higher end-of-year Math Inventory™ scores compared to students who spent less than 58 minutes in the platform.
- Among K – 6 students who identified as: Black (African American), Hispanic, and white, those who completed more lessons per week in DreamBox and those who spent more time (minutes per week) in DreamBox Math had higher spring Math Inventory™ scores.
- Among K – 6 students who had IEPs, those who completed more lessons per week in DreamBox had higher spring Math Inventory™ scores.

Conclusions

This study provides results to satisfy ESSA evidence requirements for Level II (Moderate Evidence) given the study design and positive statistically significant findings, quasi-experimental design, and large sample size.

TABLE OF CONTENTS

Introduction 3

Program Implementation Research Questions 3

Effectiveness Research Questions 3

Methods 4

Program Implementation Findings 6

Student Findings 9

Effectiveness Findings for Kindergarten – Grade 6 Students 10

Effectiveness Findings for Grades 7 – 8 Students 16

Conclusions and Recommendations 20

References 21

Appendix A. DreamBox Math Logic Model 22

Appendix B. Additional Information on Study Design and Methods 23

Appendix C. Additional Information on Program Implementation 25

Appendix D. HMH Math Inventory™ Performance Level Bands 26

Appendix E. Additional Information on K – 6 Outcome Findings 27

Appendix F. Additional Information on Grades 7 – 8 Outcome Findings 32

Introduction

DreamBox Learning recognizes that a substantial number of students in the US do not develop the mathematical proficiency needed to succeed in math beyond the middle school grades. Teacher surveys indicated that students entering algebra courses do not have sufficient prior mathematical knowledge in the areas of whole number arithmetic, fractions, and proportions (NMAP, 2008). This lack of readiness is associated with negative downstream effects for students. The DreamBox Learning K-8 Math learning solution aims to solve this problem and ensure that all students are prepared for high school mathematics, as well as STEM college and career standards.

As part of their ongoing efforts to demonstrate the efficacy of DreamBox Math, DreamBox Learning contracted with LearnPlatform, a third-party edtech research company, to examine the relationship between usage of DreamBox Math and student outcomes. After collaborating on the co-development of an updated logic model (Appendix A) for DreamBox Learning Math (Shah & Styers, 2022), LearnPlatform designed a study to satisfy ESSA Level II requirements (Moderate Evidence) with the following research questions.

Program Implementation Research Questions

1. How many DreamBox Learning Math lessons were completed by Grades K-8 students during the 2021–22 school year?
2. How much time in DreamBox Learning Math lessons was spent by Grades K-8 students during the 2021–22 school year?
3. Among DreamBox Learning Math users, what were the usage patterns?

Effectiveness Research Questions

After controlling for students' prior math achievement, grade, gender, race, ELL status, and IEP status,

4. How were different DreamBox Learning Math usage patterns related to Grade K-8 students' spring 2022 math achievement?
    a. Which usage pattern(s) of DreamBox Learning Math had the greatest impact on Grade K-8 students' spring 2022 math achievement?
5. What was the overall impact of DreamBox Learning Math on Grade K-8 students' spring 2022 math achievement?
    a. How did the impact of DreamBox Learning Math vary by student grade, and racial identity?
    b. What was the impact for students who had English language learner (ELL) status and Individualized Educational Programs (IEPs)?
6. How did end-of-year math outcomes of grades 7 – 8 who used DreamBox Math during the 2021–22 school year compare to outcomes of students who did not use the program? What was the magnitude of any observed difference?

Methods

This section of the report briefly describes the setting, participants, measures, and analysis methods.

Setting

The study included one district in Arizona and an analysis sample of K-8 students across 39 schools. Approximately half the schools (51%) are classified as serving large city locales. 44% serve large suburban locales, and 5% serve fringe rural locales.

Participants

Less than five percent of the K - 6 sample were true non-users of DreamBox Math therefore, researchers split the analytic approach; a correlative, treatment-only study was conducted for the K-6 sample and a comparative study for the grades 7-8 sample.

Grades K - 6. Students who used DreamBox Math in the 2021-22 school year (*n* = 9,555 students across 32 schools) were enrolled in kindergarten (1%), grade 1 (11%), grade 2 (15%), grade 3 (18%), grade 4 (19%), grade 5 (19%), and grade 6 (17%). In terms of demographics, the sample included the following students: 2% Asian, 3% Black, 13% Hispanic, 81% White and 1% Other; 48% identified as female. Six percent of the students had ELL status and 15% had IEPs.

Grades 7 - 8. Students in the 2021-22 school year (*n* = 4,034 students across 24 schools) were enrolled in grade 7 (52%) and grade 8 (48%). Sixty-one percent of the students were users of DreamBox Math and 39% were non-users (details about demographics differences provided in Appendix B). In terms of demographics, the sample included the following students: 2% Asian, 3% Black, 13% Hispanic, 81% White and 1% Other; 49% identified as female. Three percent of the students had ELL status and 15% had IEPs.

Measures

This study included the following measures to provide insights into DreamBox Math implementation and evidence about the potential impacts of DreamBox Math on student outcomes.

DreamBox Math Usage Metrics. Researchers utilized 2021-22 student-level usage data (i.e., average weekly lessons and weekly minutes) to inform the extent to which students used DreamBox Math during the school year and whether students' use of DreamBox Math was related to outcomes on Math Inventory™. According to DreamBox Learning, measuring intended usage of the product aligns most closely with the number of lessons students complete. DreamBox Learning recommends that students complete five lessons per week, and lesson completion is the single best indicator of student progress through the curriculum. Notably, students are credited with completing a finished lesson regardless of whether they have passed or failed it. Time may also be a practical measure of intended usage but can include a considerable amount of non-productive usage (i.e., time off task).

Standardized Student Assessments. Researchers used HMH's Math Inventory™, an adaptive, research-based assessment that reliably measures math ability and progress from Kindergarten through

Algebra II (typically taken in Grades K - 12). HMH Math Inventory assesses students' math abilities and performance based on the Quantile® Framework for Mathematics, a scientific taxonomy of more than 500 math concepts and skills that places students' readiness for math instruction and the difficulty of math tasks on the same scale (see Appendix D).

Data Analysis

Researchers used a variety of quantitative analytic approaches. First, researchers conducted descriptive statistics to describe participant characteristics and support implementation analyses. Researchers then conducted linear regressions and partial correlations to examine how DreamBox Math use related to K – 8 student math outcomes from beginning- to end-of-year. The analyses included student-level covariates to control for potential selection bias. Researchers also examined whether there were any differences between grade 7 – 8 students who used DreamBox Math during the 2021–22 school year and students who did not use the program using linear regression analysis. In addition, researchers calculated standardized effect sizes to determine the magnitude of changes in student outcomes or the standardized difference between student groups' math outcomes on HMH's Math Inventory™.

Baseline Equivalence

To ensure the validity of the study's findings and to adhere to ESSA Level II standards, the researchers assessed the equivalence of grades 7 – 8 student demographic characteristics (i.e., grade, gender, ELL status, and IEP status) and standardized assessment scores between student groups (i.e., students who used DreamBox Math and students who did not use the program). Students who used DreamBox Math were not statistically significantly different from students who did not use DreamBox Math regarding gender (effect size = 0.01), IEP status (effect size = 0.01), and beginning-of-year Math Inventory™ Quantile® scores (effect size = 0.00). However, results indicate that significantly fewer English Language Learners used DreamBox Math (effect size = -0.05). Students who used DreamBox Math were also significantly different in terms of race (effect size = 0.15) and grade (effect size = -0.15). Baseline differences with an effect size between 0.05 and 0.25 must include acceptable statistical adjustments in analyses according to What Works Clearinghouse (WWC) criteria (WWC, 2022). Therefore, ELL status, race, and grade were statistically controlled for in the final model. Researchers also included beginning-of-year Math Inventory™ Quantile® score as a covariate in the final model even though differences at baseline were not statistically significant. See Appendix B for more details regarding baseline equivalence.

Program Implementation Findings

The charts below highlight DreamBox Math use during the 2021-22 school year based on DreamBox Learning's internal usage data (Table 1). Overall, K – 6 students completed an average of 2.7 DreamBox Math lessons (SD = 2.3) per week and spent an average of 34 minutes in DreamBox Math (SD = 24) per week, while grade 7 – 8 users completed an average of 1.3 DreamBox Math lessons (SD = 1.6) per week and spent an average of 13 minutes in DreamBox Math (SD = 16) per week.

Table 1: Average DreamBox Math student usage by grade

	Number of users	Average DreamBox Math lessons completed per week	Average time (minutes) spent in DreamBox Math per week
Kindergarten	**71**	**3.2**	**33**
Grade 1	**1,074**	**3.6**	**39**
Grade 2	**1,407**	**2.7**	**36**
Grade 3	**1,716**	**2.9**	**35**
Grade 4	**1,789**	**3.2**	**39**
Grade 5	**1,843**	**2.4**	**35**
Grade 6	**1,655**	**2.1**	**27**
Total	**9,555**	**2.7**	**34**
Grade 7	**1,435**	**1.4**	**15**
Grade 8	**1,015**	**1.1**	**11**
Total	**2,450**	**1.3**	**13**

Researchers conducted a *k*-means cluster analysis to group students by similar levels of DreamBox Math usage based on the number of average weekly lessons completed and the amount of time (average weekly minutes) spent on DreamBox Math for both sets of students (i.e., K – 6 and 7 – 8 users).

For average weekly lessons, K – 6 students fell into three usage categories ranging from low usage (less than 2.4 lessons) to moderate usage (between 2.4 and 5.5 lessons), and high usage (more than 5.5 lessons) (Figure 1). Grades 7– 8 students fell into three usage categories ranging from low usage (less than 1.5 lessons) to moderate usage (between 1.5 and 4.2 lessons), and high usage (more than 4.2 lessons) (Figure 2). DreamBox Learning recommends students complete five lessons per week.

For average weekly minutes, K – 6 students fell into three usage categories ranging from low usage (less than 28 minutes) to moderate usage (between 28 and 58 minutes), and high usage (more than 58 minutes) (Figure 3). Grades 7– 8 students fell into three usage categories ranging from low usage (less than 14 minutes) to moderate usage (between 14 and 38 minutes), and high usage (more than 38 minutes) (Figure 4). Completing five lessons per week in DreamBox Math usually takes students one hour per week (see Appendix C for more details).

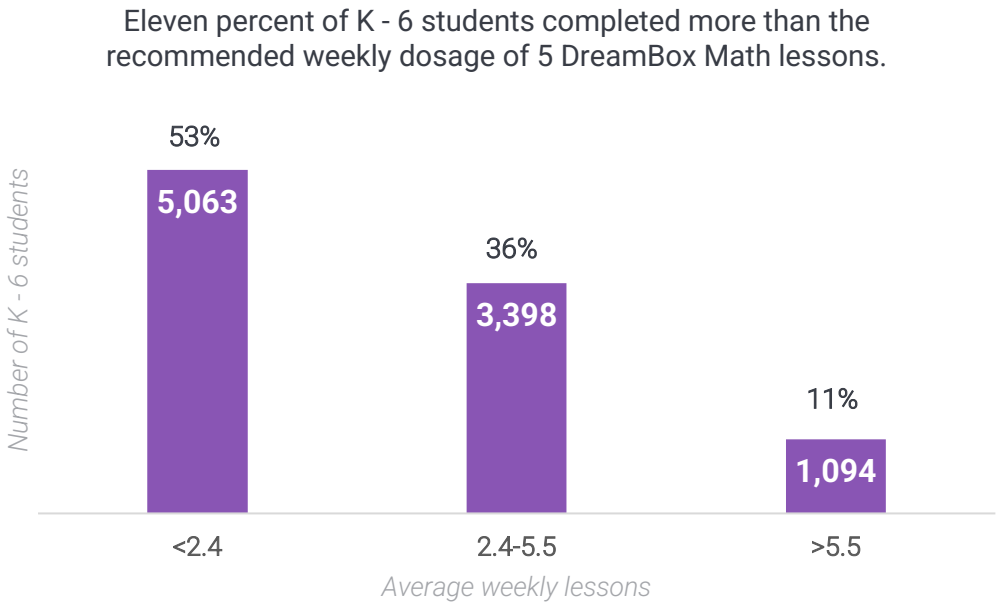


Figure 1. Overall distribution of average weekly lessons completed on DreamBox Math by K – 6 students.

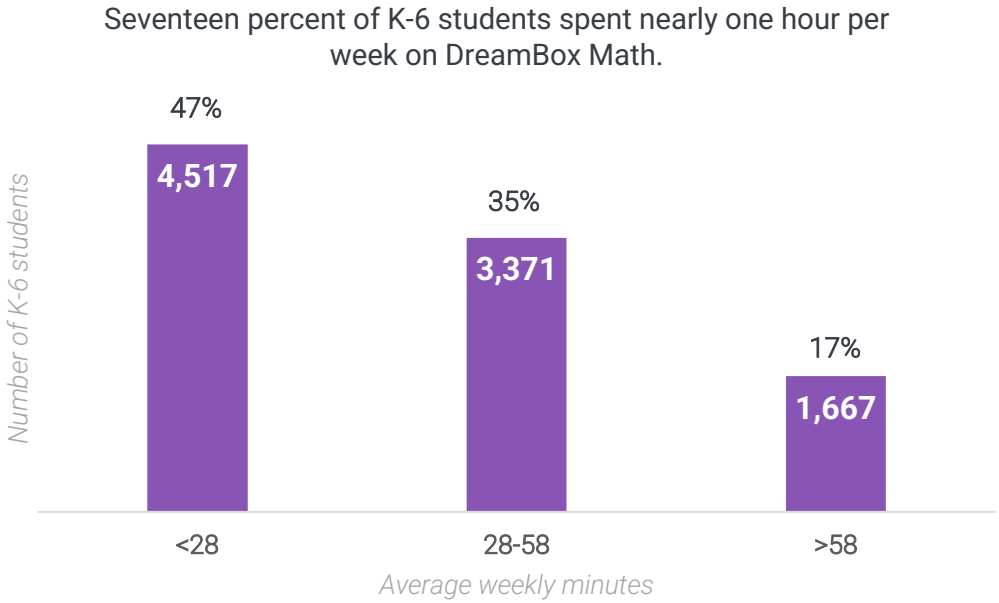


Figure 2. Overall distribution of average weekly minutes spent on DreamBox Math by K – 6 students.

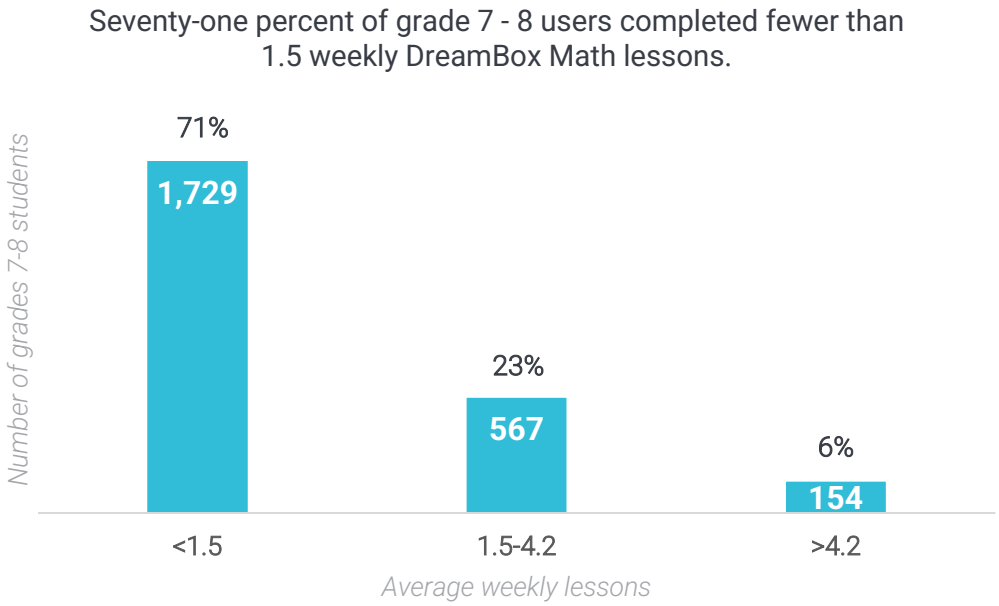


Figure 3. Overall distribution[2] of average weekly lessons completed on DreamBox Math by Grade 7 – 8 users.

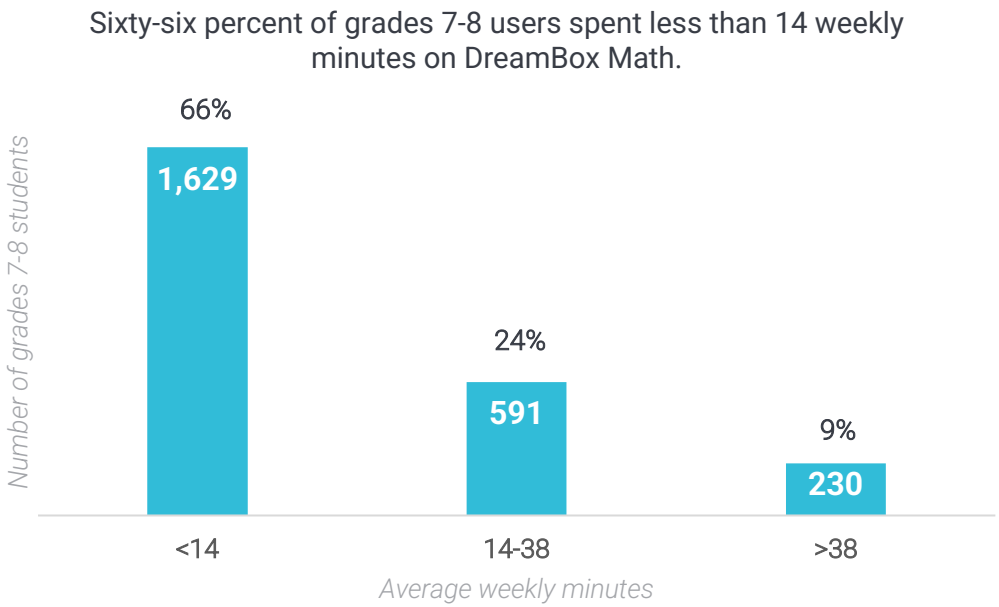


Figure 4. Overall distribution[2] of average weekly minutes spent on DreamBox Math by Grade 7 – 8 users.

[2] Percentages may not add up to one hundred due to rounding.

Student Findings

To answer the remaining study research questions, researchers conducted descriptive statistics, regression, and partial correlations. Researchers reported statistically significant findings at the p = .05 level. To determine the magnitude of the relationship, researchers calculated standardized effect sizes. Before running regression and partial correlations, researchers examined *unadjusted* Math Inventory™ Quantile® scores at the beginning and end of the year (see Figure 5). Overall, all students showed growth from beginning- to end-of-year (see Appendix D for information about HMH Math Inventory™ Performance Level Bands).

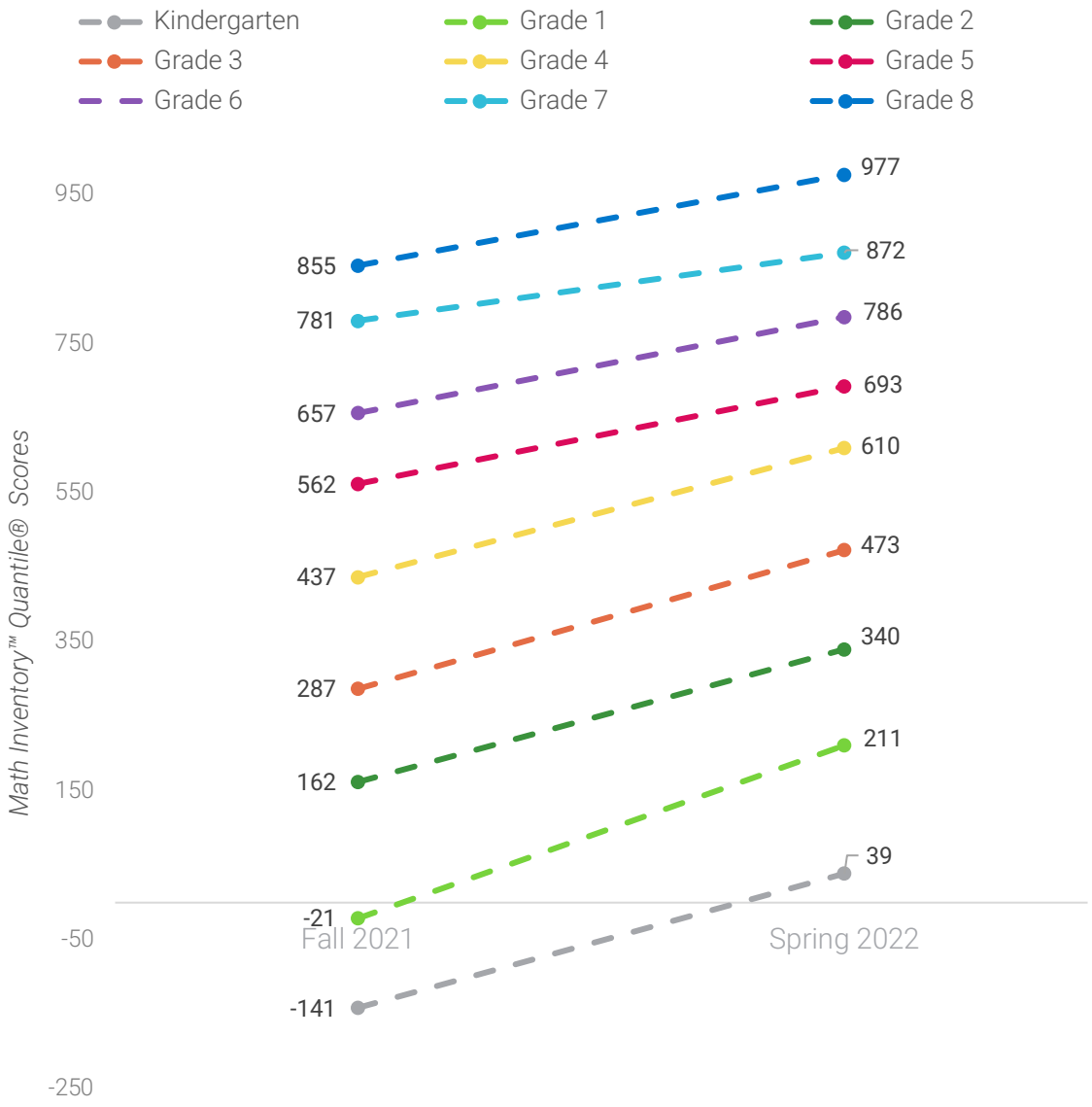


Figure 5. *Unadjusted* Math Inventory™ Quantile® scores increased from fall to spring for DreamBox Math students.

Effectiveness Findings for Kindergarten – Grade 6 Students

Researchers first examined whether greater usage of DreamBox Math related to higher spring Math Inventory™ achievement using regression models that included beginning-of-year Math Inventory™ achievement, grade, gender, race, ELL status, and IEP status as covariates. Researchers report statistically significant findings at the p = .05 level. Statistically significant findings are marked green (positive effect size) or red (negative effect size) in graphs. Findings that are not statistically significant are marked yellow (see Appendix E for details).

Overall Relationship Between Average Weekly Lessons and K – 6 Students' Outcomes on Math Inventory™

Key Finding 1. K – 6 students who completed between 2.4–5.5 weekly lessons (moderate use) and between 5.5–15.7 weekly lessons (high use) in DreamBox Math had higher end-of-year Math Inventory™ Quantile® scores compared to students who completed less than 2.4 weekly lessons (low use). Further, K – 6 students who completed between 5.5–15.7 weekly lessons (high use) in DreamBox Math had higher end-of-year Math Inventory™ Quantile® scores compared to students who completed between 2.4–5.5 weekly lessons (moderate use) (Figures 6).

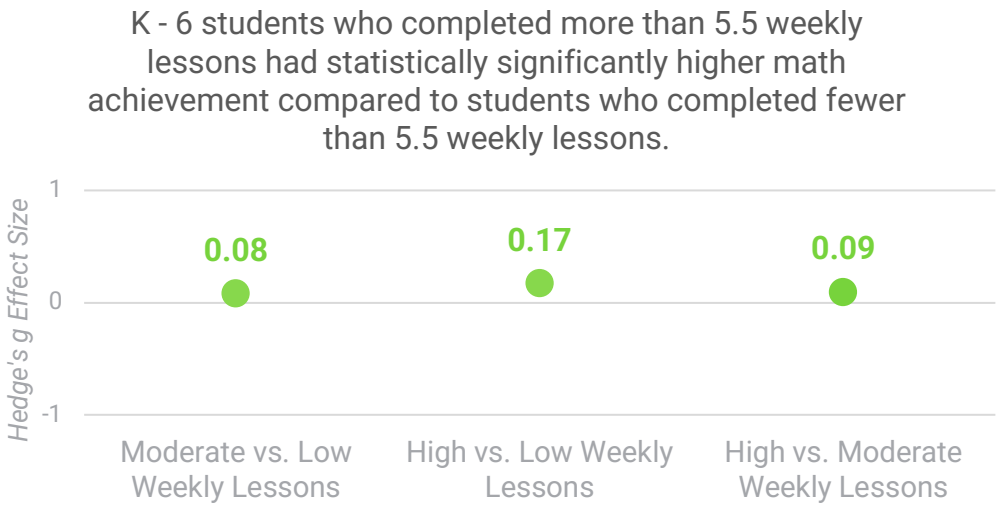


Figure 6. Relationship between average weekly lessons and K – 6 students' outcomes on Math Inventory™ in terms of Hedge's *g* effect sizes

Overall Relationship Between Average Weekly Minutes and K – 6 Students' Outcomes on Math Inventory™

Key Finding 2. K – 6 students who spent between 58–120 weekly minutes (high use) in DreamBox Math had higher end-of-year Math Inventory™ Quantile® scores compared to students who spent less than 58 weekly minutes (low and moderate use) in the platform (Figures 7).

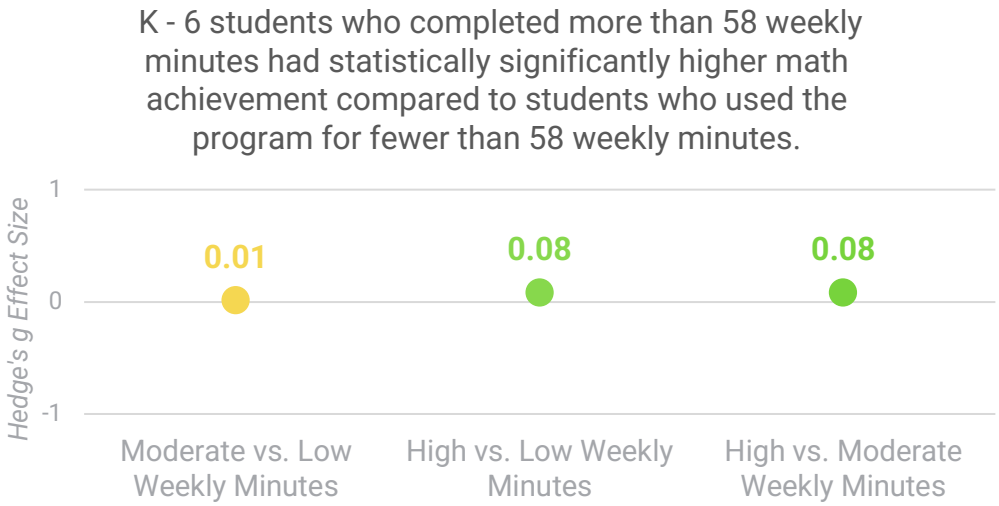


Figure 7. Relationship between average weekly minutes and student math outcomes on Math Inventory™ in terms of Hedge's *g* effect sizes

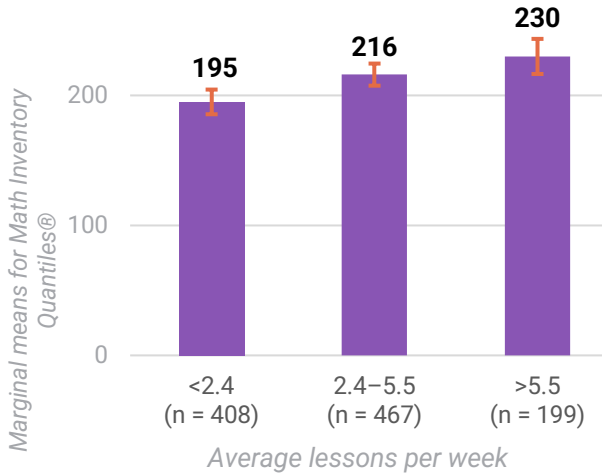
Relationship Between Average Weekly Lessons and Minutes and K – 6 Students' Outcomes on Math Inventory™ by Grade

Researchers examined whether greater usage of DreamBox Math related to higher spring Math Inventory™ achievement for each grade. To allow for better interpretability of results, marginal means charts for each grade that had statistically significant results i.e., Grade 1 – 6, are presented below. The orange vertical lines at the top of each bar represent a 95% confidence interval.

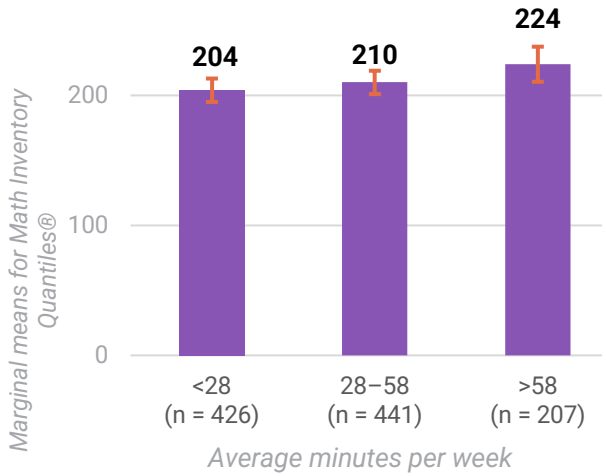
Key Finding 3. Positive results were observed for all grade levels except Grade 6, and the differences for Grades 1 – 5 students were statistically significant relationships ($p < .05$), such that students who completed more DreamBox Math weekly lessons had higher spring Math Inventory™ scores.

Key Finding 4. Among Grades 1, 3, 4 and 5 students, there was a positive, statistically significant relationship ($p < .05$), such that students who spent more time in DreamBox Math (weekly minutes) had higher spring Math Inventory™ scores. The results for kindergarten were positive but not statistically significant. The results for Grade 2 were mixed; there were statistically significant positive and negative results. Finally, the results for Grade 6 were mostly negative and with one statistically significant negative result. Details about *p*-values and effect sizes are given in Appendix E.

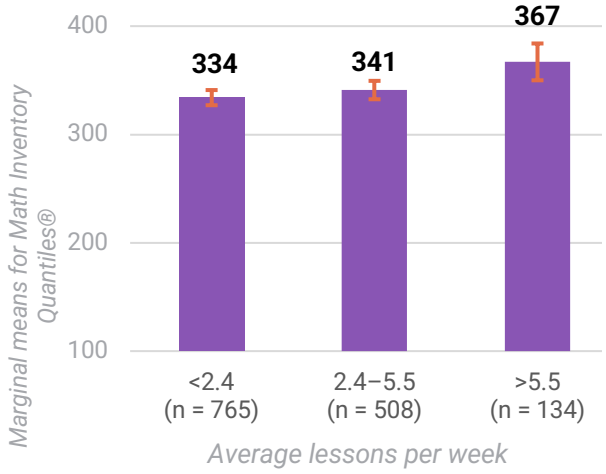
Grade 1 students who completed more than 2.4 *weekly lessons* (high and moderate use) had significantly higher scores than students who completed fewer than 2.4 lessons (low use).



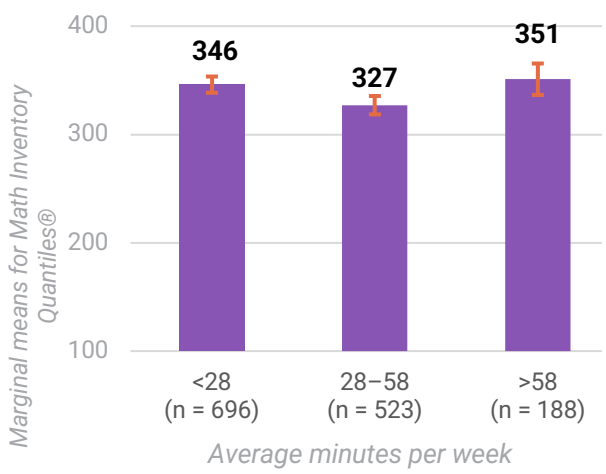
Grade 1 students who spent more than 58 *weekly minutes* on DreamBox Math (high use) had significantly higher scores than students who spent less than 28 weekly minutes (low use).



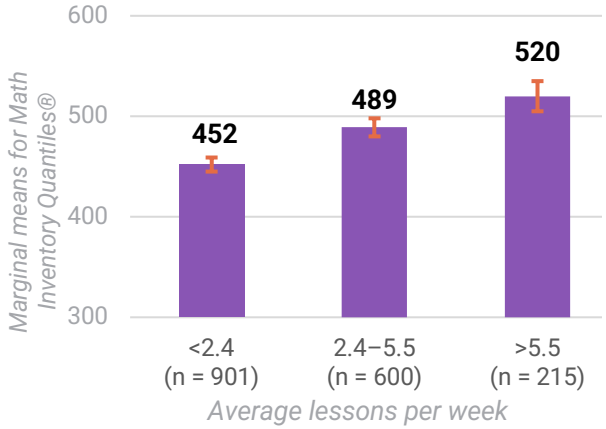
Grade 2 students who completed more than 5.5 weekly lessons (high use) had significantly higher scores than students who completed fewer than 5.5 lessons (low and moderate use).



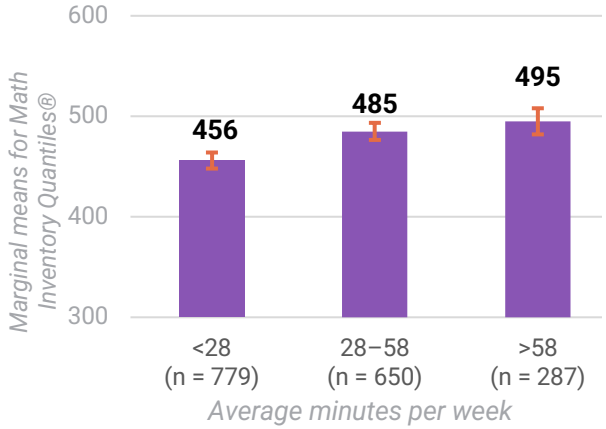
Grade 2 students who spent more than 58 weekly minutes on DreamBox Math (high use) or less than 28 minutes (low use) had significantly higher scores than students who spent between 28 and 58 weekly minutes (low use).



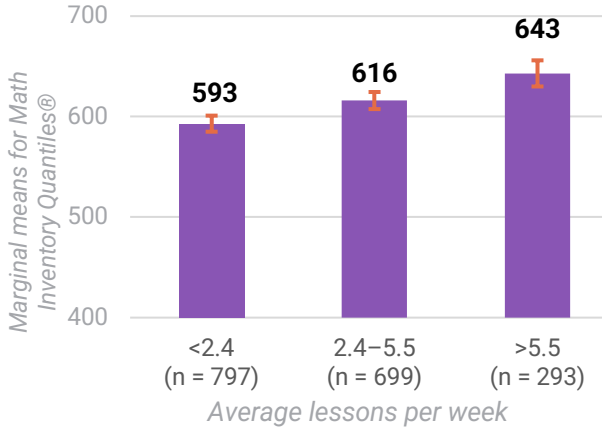
Grade 3 students who completed more than 2.4 weekly lessons (high and moderate use) had significantly higher scores than students who completed fewer than 2.4 lessons (low use). Students who completed more than 5.5 lessons (high use) had significantly higher scores than students who completed between 2.4 and 5.5. lessons (moderate use).



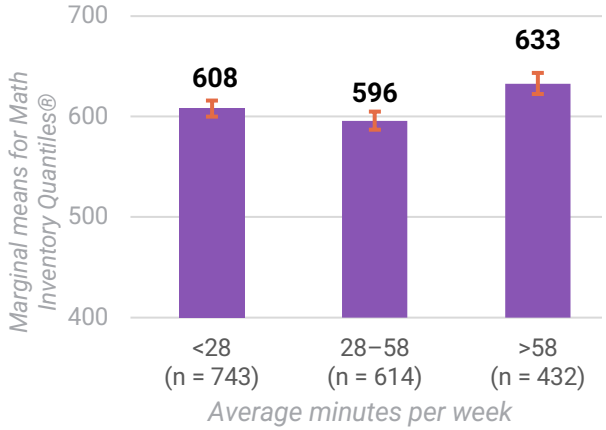
Grade 3 students who spent more than 28 weekly minutes on DreamBox Math (high and moderate use) had significantly higher scores than students who spent fewer than 28 weekly minutes (low use).



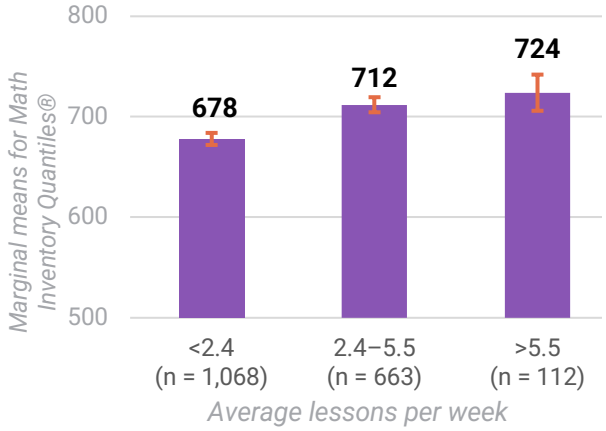
Grade 4 students who completed more than 2.4 weekly lessons (high and moderate use) had significantly higher scores than students who completed fewer than 2.4 lessons (low use). Students who completed more than 5.5 lessons (high use) had significantly higher scores than students who completed between 2.4 and 5.5. lessons (moderate use).



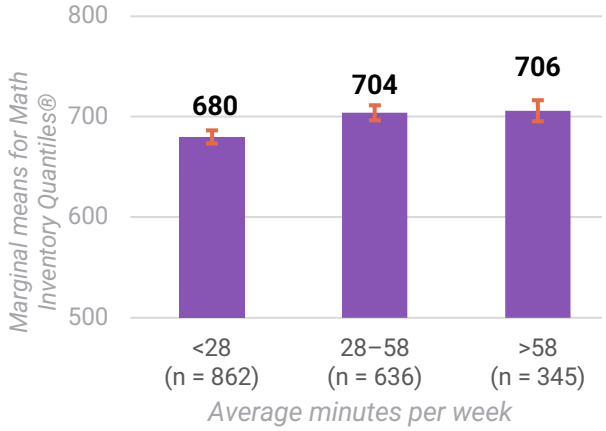
Grade 4 students who spent more than 58 weekly minutes on DreamBox Math (high) had significantly higher scores than students who spent fewer than 58 weekly minutes (low and moderate use).



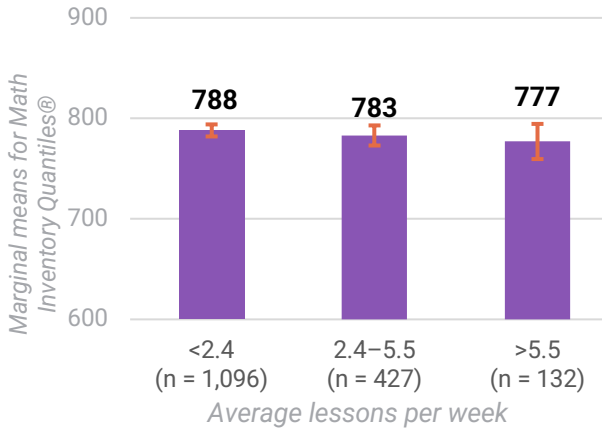
Grade 5 students who completed more than 2.4 weekly lessons (high and moderate use) had significantly higher scores than students who completed fewer than 2.4 lessons (low use).



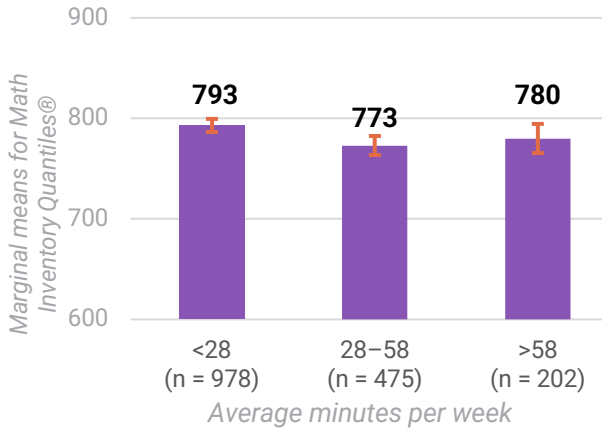
Grade 5 students who spent more than 28 weekly minutes on DreamBox Math (high and moderate use) had significantly higher scores than students who spent fewer than 28 weekly minutes (low use).



There were no statistically significant relationships between weekly lessons completed and math outcomes among Grade 6 students.



Grade 6 students who spent fewer than 28 weekly minutes on DreamBox Math (low use) had significantly higher scores than students who spent between 28 and 58 weekly minutes (moderate use).



Relationship Between DreamBox Usage and K – 6 Students' Outcomes on Math Inventory™ by Race

Key Finding 5. Among students who identified as Black (African American), students who identified as Hispanic, and students who identified as white, there was a positive, statistically significant relationship such that students who completed more lessons in DreamBox (lessons weekly) had higher spring Math Inventory™ scores (*r* values = 0.20, 0.22, and 0.12; *p* values < .05, respectively) and those who spent more time in DreamBox Math (weekly minutes) also had higher spring Math Inventory™ scores (*r* values = 0.16, 0.16, and 0.04; *p* values < .05, respectively) (Appendix E).

Relationship Between DreamBox Usage and K – 6 Students' Outcomes on Math Inventory™ by English Language Learner (ELL) Status

Key Finding 6. Among students who had ELL status, there was a positive, statistically significant relationship ($p < .05$), such that students who completed more lessons in DreamBox (weekly lessons) and those who spent more time in DreamBox Math (weekly minutes) had higher spring Math Inventory™ scores (r values = 0.19 and 0.13, respectively; Appendix E).

Relationship Between DreamBox Usage and K – 6 Students' Outcomes on Math Inventory™ by Individualized Educational Program (IEP)

Key Finding 7. Among students who had an IEP, there was a positive, statistically significant relationship ($p < .05$), such that students who completed more lessons in DreamBox (weekly lessons) and those who spent more time in DreamBox Math (weekly minutes) had higher spring Math Inventory™ scores (r values = 0.10 and 0.07, respectively; Appendix E).

Effectiveness Findings for Grades 7 – 8 Students

Researchers first examined whether greater usage of DreamBox Math related to higher spring Math Inventory™ achievement using regression models that included beginning-of-year Math Inventory™ achievement, grade, gender, race, ELL status, and IEP status as covariates. Statistically significant findings are marked green (positive effect size) or red (negative effect size) in graphs. Findings that are not statistically significant are marked yellow. Researchers also examined whether there were any differences between students who used DreamBox Math during the 2021–22 school year and students who did not use the program using linear regression analysis. Of note, Appendix F provides additional information on these analyses and findings, and researchers reported statistically significant findings at the p = .05 level. Standardized effect sizes were also calculated to assist with interpretation.

Overall Relationship Between Average Weekly Lessons and Grades 7 – 8 Students' Outcomes on Math Inventory™

Key Finding 8. Grades 7 – 8 students who completed more than 1.5 weekly lessons (high and moderate use) had higher end-of-year Math Inventory™ Quantile® scores compared to students who completed fewer than 1.5 weekly lessons (low and moderate use). Further, grades 7 – 8 students who completed more than 4.2 weekly lessons (high use) had higher end-of-year Math Inventory™ Quantile® scores compared to students who completed between 1.5 and 4.2 weekly lessons (moderate use) (Figures 8).

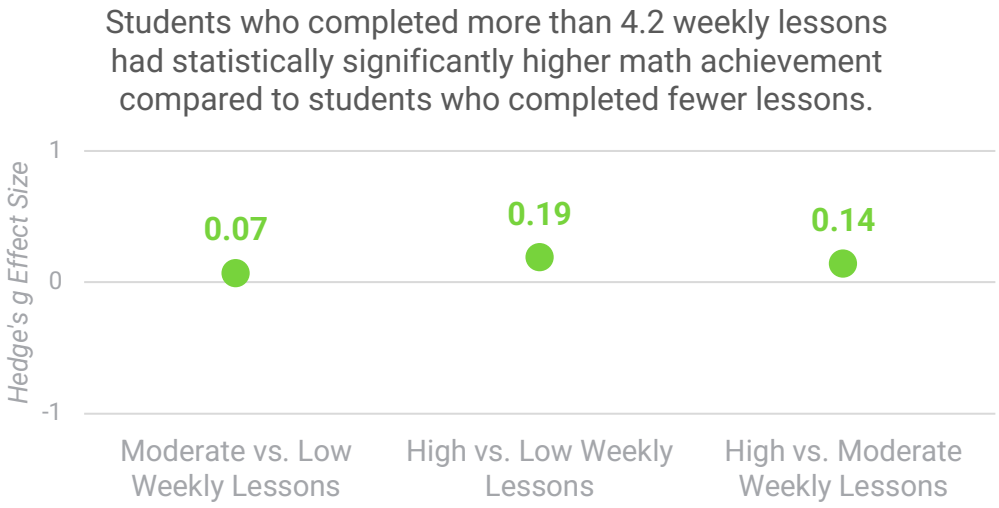


Figure 8. Relationship between average weekly lessons and grades 7 – 8 students' outcomes on Math Inventory™ in terms of Hedge's *g* effect sizes

Overall Relationship Between Average Weekly Minutes and Grades 7 – 8 Students' Outcomes on Math Inventory™

Key Finding 9. Students who spent greater than 38 weekly minutes (high use) in DreamBox Math had higher end-of-year Math Inventory™ Quantile® scores compared to students who spent less than 38 weekly minutes (low and moderate use) in the platform (Figures 9).

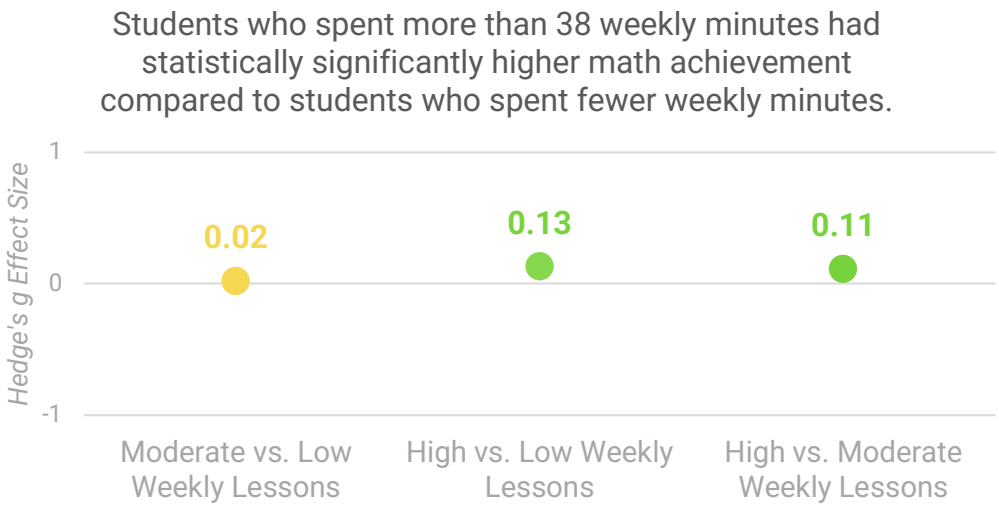


Figure 9. Relationship between average weekly minutes and grades 7 – 8 students' outcomes on Math Inventory™ in terms of Hedge's *g* effect sizes

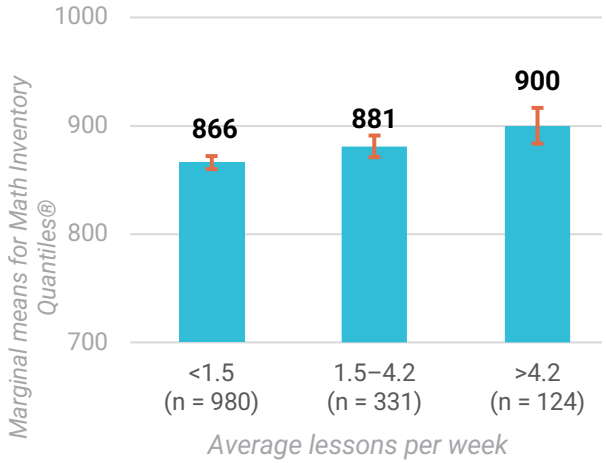
Relationship Between Average Weekly Lessons and Minutes and Grades 7 – 8 Students' Outcomes on Math Inventory™ by Grade

Researchers examined whether greater usage of DreamBox Math related to higher spring Math Inventory™ achievement for each grade. To allow for better interpretability of results, marginal means charts are presented below. The orange vertical lines at the top of each bar represent a 95% confidence interval.

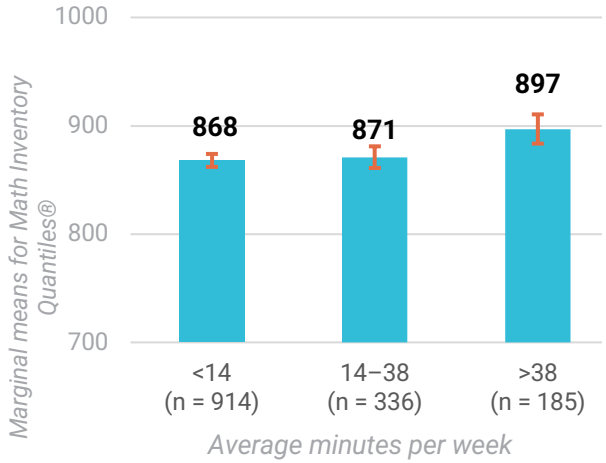
Key Finding 10. Positive results were observed for grade 7, and the differences were statistically significant relationships ($p < .05$), such that grade 7 students who completed more DreamBox Math weekly lessons had higher spring Math Inventory™ scores. While the grade 8 results were positive, none of the relationships were statistically significant.

Key Finding 11. Among grade 7 students, there was a positive, statistically significant relationship ($p < .05$), such that students who spent more time in DreamBox Math (weekly minutes) had higher spring Math Inventory™ scores. While the grade 8 results were mixed, none of the relationships were statistically significant. Details about *p*-values and effect sizes are given in Appendix F.

Grade 7 students who completed more than 1.5 weekly lessons (high and moderate use) had statistically significant higher scores than students who completed fewer than 1.5 lessons (low use).



Grade 7 students who spent more than 38 weekly minutes on DreamBox Math (high use) had statistically significant higher scores than students who spent less than 38 weekly minutes (low and moderate use).



Relationship Between DreamBox Usage and Grades 7 – 8 Students' Outcomes on Math Inventory™ by Race

Key Finding 12. Among grades 7 – 8 students who identified as white, there was a positive, statistically significant relationship such that students who completed more lessons in DreamBox (lessons weekly) had higher spring Math Inventory™ scores ($r = 0.09$; $p < .05$) and those who spent more time in DreamBox Math (weekly minutes) also had higher spring Math Inventory™ achievement ($r = 0.05$; $p < .05$) (Appendix F).

Relationship Between DreamBox Usage and Grades 7 – 8 Students' Outcomes on Math Inventory™ by English Language Learner (ELL) Status

Key Finding 13. Among students who had ELL status, there was a positive, statistically significant relationship ($p < .05$), such that students who completed more lessons in DreamBox (weekly lessons) and those who spent more time in DreamBox Math (weekly minutes) had higher spring Math Inventory™ scores (r values = 0.37 and 0.32, respectively; Appendix F).

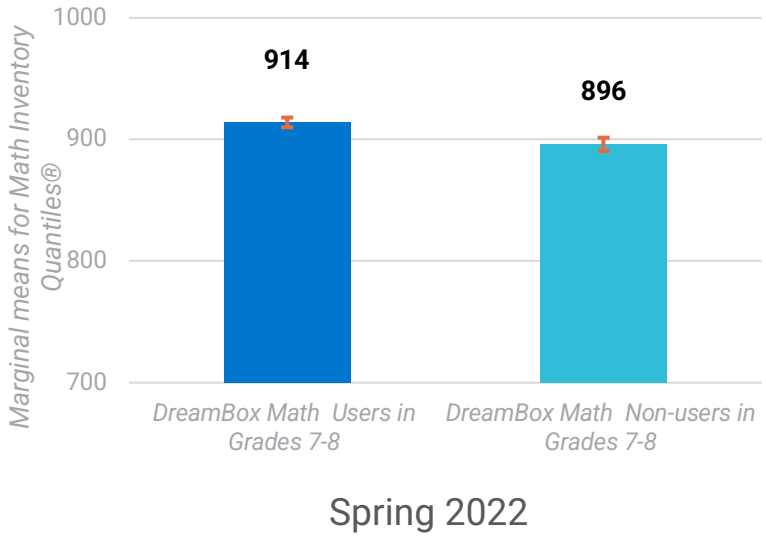
Relationship Between DreamBox Usage and Grades 7 – 8 Students' Outcomes on Math Inventory™ by Individualized Educational Program (IEP)

Key Finding 14. Among students who had an IEP, there were no statistically significant relationships.

Differences in End-of-Year Math Outcomes Between Grades 7 – 8 Students who used DreamBox Math and Students Who Did Not Use the Program

To determine whether there were differences between students who used DreamBox Math and students who did not use the program, researchers conducted linear regression analysis to examine whether DreamBox usage predicted differences in Math Inventory™ end-of-the-year Quantile® scores with beginning-of-the-year Quantile® scores, grade, race, ELL status, and propensity score weights as covariates. Results show that students who used DreamBox Math scored statistically significantly higher on the end-of-the-year Math Inventory™ when compared to students who did not use the program (see Key Finding 15).

Key Finding 15. Grades 7 – 8 students who used DreamBox Math had higher scores on Math Inventory™ at the end of the year than students who did not use the program (effect size = 0.12; $p < 0.001$).



Note: The orange vertical lines at the top of each bar represent a 95% confidence interval.

Conclusions and Recommendations

In sum, the findings support a relationship between DreamBox Math usage and improved math skills for K-8 students. Further, given the positive outcome findings of the impact analysis among the grades 7 – 8 sample, this study provides results to satisfy ESSA evidence requirements for Level II (Moderate Evidence). Specifically, this study met the following criteria:

✔ Quasi-experimental design

✔ Proper design and implementation

✔ Statistical controls through covariates

✔ At least one statistically significant, positive finding

Researchers recommend the following next steps:

- K – 6 students who completed at least 5.5 weekly lessons and grade 7 – 8 users who completed at least 4.2 weekly lessons had higher scores on Math Inventory™. DreamBox Learning should continue to explore ideal implementation at various grade-bands among other sites using this baseline information.
- DreamBox Learning should consider recruiting a comparison district for K – 6 students to better understand how elementary school students who use DreamBox Math compare to elementary school students using other math programs.

Acknowledgements

The authors would like to extend their deepest thanks to Avery Wall who supported the preparation of this report in numerous ways.

References

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Washington, DC: National Academies Press.

What Works Clearinghouse. (2022). What Works Clearinghouse procedures and standards handbook, version 5.0. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). This report is available on the What Works Clearinghouse website at https://ies.ed.gov/ncee/wwc/Handbooks

Appendix A. DreamBox Math Logic Model

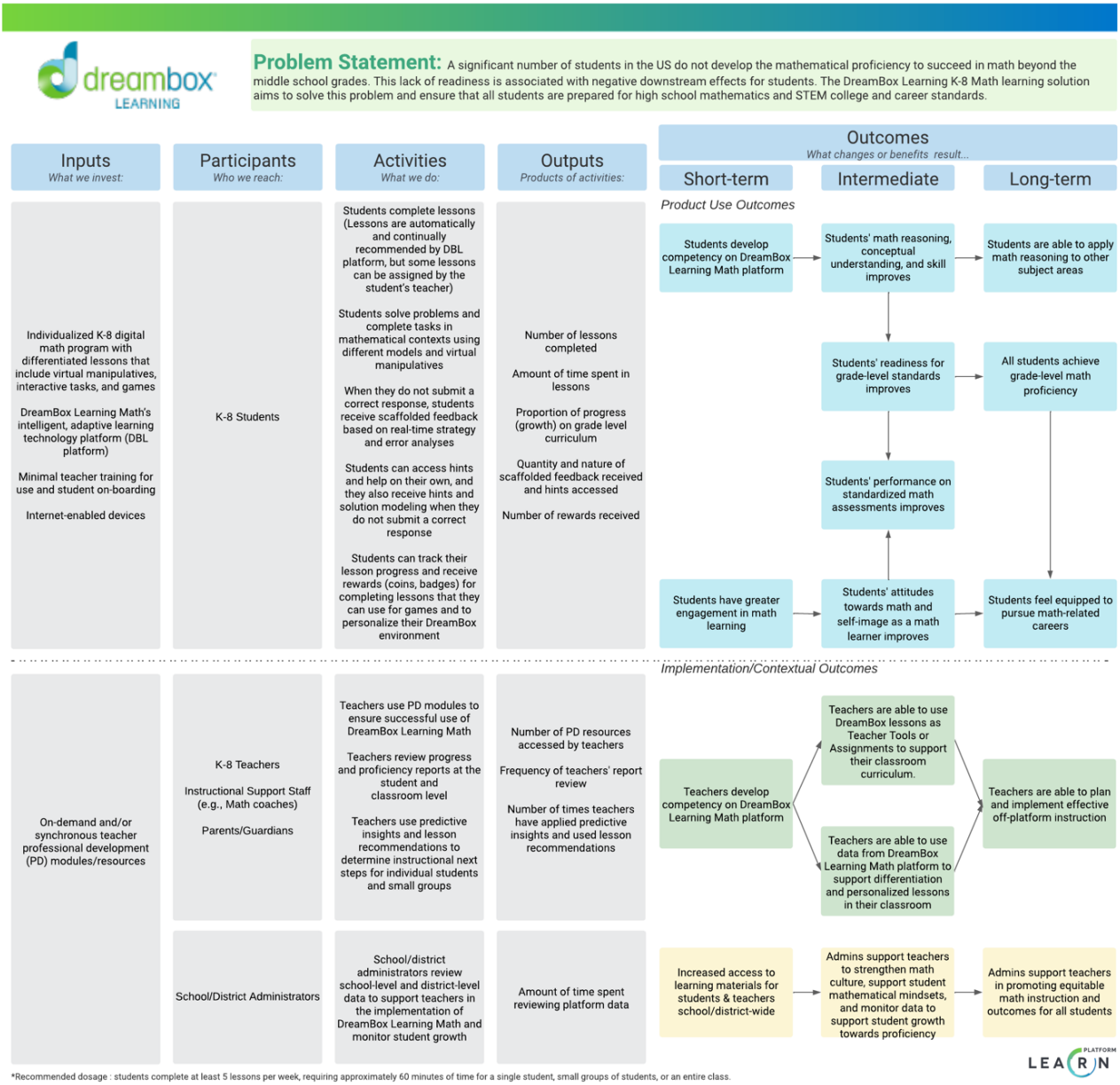


Figure 1. DreamBox Learning Math logic model

Prepared for DreamBox Learning, September 2022

Appendix B. Additional Information on Study Design and Methods

Propensity Score Weighting

To help make the grade 7 – 8 student groups (i.e., students who used DreamBox Math and students who did not use DreamBox Math) as comparable as possible, propensity score weights were calculated for each student. To calculate propensity scores, researchers conducted binary logistic regression with student group as the dependent variable and grade, gender, ELL and IEP status, and Math Inventory™ beginning-of-year Quantile® scores as the covariates. The probability was saved as a new variable. Weights were calculated by dividing one by the probability (one/probability). Students without a weight were dropped from the final analytic sample. All analyses that included students who did not use DreamBox Math included these weights.

Baseline Equivalence

Researchers conducted baseline equivalence analyses to determine whether there were baseline differences in characteristics between students who used MMA and students who did not use the program during the 2021–22 school year. Specifically, researchers used chi-square analyses on student-level demographics and linear regressions on beginning-of-year Quantile® scores.

As noted in Table B1, there were no statistically significant differences between groups regarding gender and IEP status. However, students who used DreamBox Math were statistically significantly different regarding ELL status, race, and grade.

Table B1. Baseline Equivalence Analysis of Grade 7-8 Student-Level Demographics by User Group

Characteristics	Users (*n* = 2,450)		Non-users (*n* =1,584)		Chi-squared	*p*-Value	Effect Size
	Percent	*N*	Percent	*N*			
Gender							
Male	52	1,272	51	801	0.70	.40	0.01
Female	48	1,178	49	783			
Individualized Educational Program (IEP)							
Yes	15	363	15	244	0.26	.61	-0.01
No	85	2,087	85	1,340			
English Language Learner (ELL) Status							
Yes	2	47	4	58	11.53	<.001***	-0.05
No	98	2,403	96	1,526			

Characteristics	Users (*n* = 2,450)		Non-users (*n* =1,584)		Chi-squared	*p*-Value	Effect Size
	Percent	*N*	Percent	*N*			
Race							
Asian	2	52	3	45	85.83	<.001***	0.15
Black	2	54	4	58			
Hispanic	9	225	18	283			
White	85	2,093	74	1,173			
Grade							
7	68	1,435	32	671	101.32	<.001***	-0.16
8	53	1,015	47	913			

+ *p*<0.1, * *p*<0.05, ** *p*<0.01, *** *p*<0.001

As presented in Table B2, results from the weighted regression analysis reveal that there was no statistically significant difference between student groups, with a Hedge's *g* effect size indicating that researchers need not include beginning-of-year Quantile® scores in the final model.

Table B2. Baseline Equivalence Analysis of Grade 7-8 Quantile® Beginning-of-Year Scores by User Group

Outcome Variable	Coefficient	Standard Error	*t*-value	*p*-value	Effect Size
Quantile® beginning-of-year scores	-0.19	4.77	-0.04	.97	-0.001

+ *p*<0.1, * *p*<0.05, ** *p*<0.01, *** *p*<0.001

Appendix C. Additional Information on Program Implementation

Table C1. Descriptive statistics for the weekly lessons' usage categories for K – 6 sample

Usage categories: weekly lessons	*n*	Mean	SD
<2.4 weekly lessons	5,063	1.1	0.7
2.4–5.5 weekly lessons	3,398	3.7	0.9
>5.5 weekly lessons	1,094	7.4	1.7

Table C2. Descriptive statistics for the weekly minutes' usage categories for K – 6 sample

Usage categories: weekly minutes	n	Mean	SD
<28 weekly minutes	4,517	14	8
28–58 weekly minutes	3,371	42	9
>58 weekly minutes	1,667	75	14

Table C3. Descriptive statistics for the weekly lessons' usage categories for grades 7 – 8 sample

Usage categories: weekly lessons	*n*	Mean	SD
<1.5 weekly lessons	1,729	0.4	0.4
1.5–4.2 weekly lessons	567	2.6	0.7
>4.2 weekly lessons	154	5.8	1.6

Table C4. Descriptive statistics for the weekly minutes' usage categories for grades 7 – 8 sample

Usage categories: weekly minutes	*n*	Mean	SD
<14 weekly minutes	1,629	4	4
14–38 weekly minutes	591	24	7
>38 weekly minutes	230	51	13

Appendix D. HMH Math Inventory™ Performance Level Bands

Math Inventory™ reports categorize student performance in Quantiles® into bands, these are reproduced from the FAQ page linked here.

Grade	Below Basic	Basic	Proficient	Advanced
K	EM244 – EM75	EM74 – 8	9 – 117	118 – 295+
1	EM2435– EM15	16 – 116	117 – 232	233 – 384+
2	EM233 – 141	142 – 270	271 – 382	383 – 600+
3	EM151 – 276	277 – 380	381 – 545	546 – 815+
4	EM110 – 389	390 – 533	534 – 629	630 – 929+
5	77 – 539	540 – 644	645 – 771	772 – 1,045+
6	125 – 659	660 – 784	785 – 890	891 – 1,138+
7	393 – 752	753 – 880	881 – 970	971 – 1,141+
8	422 – 845	846 – 1,000	1,001 – 1,089	1,090 – 1,296+

Appendix E. Additional Information on K – 6 Outcome Findings

Overall Relationship Between Average Weekly Lessons and K – 6 Students' Outcomes on Math Inventory™

Table E1. Math Outcomes on Math Inventory™ by Weekly Lessons on DreamBox Math (covariates: beginning-of-year (BOY) Math Inventory™ score, grade, gender, race, ELL status, and IEP status)

Group Comparisons	Coefficient	Standard Error	t-value	p>|t|	Effect Size
<2.4 weekly lessons compared to 2.4–5.5 weekly lessons	21.45	2.34	9.17	< .001***	0.08
<2.4 weekly lessons compared to >5.5 weekly lessons	42.31	3.53	11.98	< .001***	0.17
2.4–5.5 weekly lessons compared to >5.5 weekly lessons	20.86	3.64	5.73	< .001***	0.09

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Overall Relationship Between Average Weekly Minutes and K – 6 Students' Outcomes on Math Inventory™

Table E2. Math Outcomes on Math Inventory™ by Weekly Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, grade, gender, race, ESL status, and IEP status)

Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<28 weekly minutes compared to 28 – 58 weekly minutes	2.46	2.41	1.02	.307	0.01
<28 weekly minutes compared to >58 weekly minutes	21.07	3.02	6.97	< .001***	0.08
28 – 58 weekly minutes compared to >58 weekly minutes	18.61	3.15	5.91	< .001***	0.08

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Relationship Between Average Weekly Lessons and K – 6 Students' Outcomes on Math Inventory™ by Grade

Table E3. Kindergarten Math Outcomes on Math Inventory™ by Weekly Lessons and Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, gender, race, and ELL and IEP status)

Kindergarten Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<2.4 weekly lessons compared to 2.4–5.5 weekly lessons	15.82	19.40	0.82	.418	0.17
<2.4 weekly lessons compared to >5.5 weekly lessons	43.34	29.37	1.48	.145	0.40
2.4–5.5 weekly lessons compared to >5.5 weekly lessons	27.51	27.97	0.98	.329	0.36

Kindergarten Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<28 weekly minutes compared to 28 – 58 weekly minutes	0.20	20.72	0.01	.992	0.00
<28 weekly minutes compared to >58 weekly minutes	44.65	33.16	1.35	.183	0.39
28 – 58 weekly minutes compared to >58 weekly minutes	44.45	34.58	1.29	.203	0.63

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table E4. Grade 1 Math Outcomes on Math Inventory™ by Weekly Lessons and Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, gender, race, and ELL and IEP status)

Grade1 Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<2.4 weekly lessons compared to 2.4–5.5 weekly lessons	20.31	6.53	3.11	.002**	0.15
<2.4 weekly lessons compared to >5.5 weekly lessons	34.91	8.52	4.1	< .001***	0.26
2.4–5.5 weekly lessons compared to >5.5 weekly lessons	14.60	8.30	1.76	.079+	0.12
<28 weekly minutes compared to 28 – 58 weekly minutes	6.11	6.59	0.93	.354	0.05
<28 weekly minutes compared to >58 weekly minutes	19.57	8.32	2.35	.019*	0.15
28 – 58 weekly minutes compared to >58 weekly minutes	13.46	8.23	1.64	.102	0.11

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table E5. Grade 2 Math Outcomes on Math Inventory™ by Weekly Lessons and Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, gender, race, and ELL and IEP status)

Grade 2 Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<2.4 weekly lessons compared to 2.4–5.5 weekly lessons	6.92	5.77	1.20	.231	0.05
<2.4 weekly lessons compared to >5.5 weekly lessons	32.43	9.44	3.44	<.001**	0.21
2.4–5.5 weekly lessons compared to >5.5 weekly lessons	25.51	9.71	2.63	.009**	0.18
<28 weekly minutes compared to 28 – 58 weekly minutes	-19.43	5.80	-3.35	<.001**	-0.13
<28 weekly minutes compared to >58 weekly minutes	5.13	8.25	0.62	.534	0.04
28 – 58 weekly minutes compared to >58 weekly minutes	24.57	8.49	2.89	.004**	0.16

Table E6. **Grade 3** Math Outcomes on Math Inventory™ by Weekly Lessons and Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, gender, race, and ELL and IEP status)

Grade 3 Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<2.4 weekly lessons compared to 2.4–5.5 weekly lessons	37.87	5.83	6.49	< .001***	0.23
<2.4 weekly lessons compared to >5.5 weekly lessons	67.97	8.46	8.03	< .001***	0.41
2.4–5.5 weekly lessons compared to >5.5 weekly lessons	30.10	8.71	3.45	<.001***	0.20
<28 weekly minutes compared to 28 – 58 weekly minutes	28.89	5.91	4.89	<.001***	0.17
<28 weekly minutes compared to >58 weekly minutes	38.71	7.68	5.04	< .001***	0.23
28 – 58 weekly minutes compared to >58 weekly minutes	9.82	7.84	1.25	.211	0.06

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table E7. **Grade 4** Math Outcomes on Math Inventory™ by Weekly Lessons and Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, gender, race, and ELL and IEP status)

Grade 4 Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<2.4 weekly lessons compared to 2.4–5.5 weekly lessons	23.05	5.80	3.97	< .001***	0.13
<2.4 weekly lessons compared to >5.5 weekly lessons	49.86	7.67	6.5	< .001***	0.27
2.4–5.5 weekly lessons compared to >5.5 weekly lessons	26.81	7.76	3.45	<.001***	0.17
<28 weekly minutes compared to 28 – 58 weekly minutes	-11.64	6.13	-1.90	.058+	-0.06
<28 weekly minutes compared to >58 weekly minutes	25.55	6.79	3.76	< .001***	0.14
28 – 58 weekly minutes compared to >58 weekly minutes	37.18	7.03	5.29	< .001***	0.22

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table E8. **Grade 5** Math Outcomes on Math Inventory™ by Weekly Lessons and Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, gender, race, and ELL and IEP status)

Grade 5 Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<2.4 weekly lessons compared to 2.4–5.5 weekly lessons	33.68	4.81	7.01	< .001***	0.21
<2.4 weekly lessons compared to >5.5 weekly lessons	45.92	9.61	4.78	< .001***	0.27

Grade 5 Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
2.4–5.5 weekly lessons compared to >5.5 weekly lessons	12.24	9.86	1.24	.214	0.09
<28 weekly minutes compared to 28 – 58 weekly minutes	23.28	5.12	4.54	< .001***	0.14
<28 weekly minutes compared to >58 weekly minutes	25.94	6.20	4.18	< .001***	0.15
28 – 58 weekly minutes compared to >58 weekly minutes	2.66	6.50	0.41	.683	0.02

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table E9. Grade 6 Math Outcomes on Math Inventory™ by Weekly Lessons and Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, gender, race, and ELL and IEP status)

Grade 6 Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<2.4 weekly lessons compared to 2.4–5.5 weekly lessons	-4.89	5.89	-0.83	.407	-0.03
<2.4 weekly lessons compared to >5.5 weekly lessons	-10.50	9.57	-1.1	.273	-0.06
2.4–5.5 weekly lessons compared to >5.5 weekly lessons	-5.61	10.31	-0.54	.586	-0.03
<28 weekly minutes compared to 28 – 58 weekly minutes	-20.47	5.75	-3.56	< .001***	-0.12
<28 weekly minutes compared to >58 weekly minutes	-13.33	8.02	-1.66	.097+	-0.08
28 – 58 weekly minutes compared to >58 weekly minutes	7.14	8.70	0.82	.412	0.04

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Relationship Between Average Weekly Lessons and K – 6 Students' Outcomes on Math Inventory™ by Race, ELL Status, IEP Status

Table E10. Partial correlation coefficients and associated *p*-values for analyses by race, ELL status, and IEP status (covariates: grade, gender, BOY Math Inventory™ score)

	n	Partial Correlation Coefficient	*p*-value
Asian	215	0.08	.273
Black/African American	316	0.20	<.001***
Hispanic	1,230	0.22	<.001***
White	7,693	0.12	<.001***

ELL Status	552	0.19	<.001***
IEP Status	1,450	0.10	<.001***

+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Relationship Between Average Weekly Minutes and K – 6 Students' Outcomes on Math Inventory™ by Race, FRL, and ELL Status

Table E11. Partial correlation coefficients and associated p-values for analyses by race, ELL status, and IEP status (covariates: grade, gender, BOY Math Inventory™ score)

	n	Partial Correlation Coefficient	*p*-value
Asian	215	-0.02	.818
Black/African American	316	0.16	.004**
Hispanic	1,230	0.16	<.001***
White	7,693	0.04	<.001***
ELL Status	552	0.13	.003**
IEP Status	1,450	0.07	<.012*

+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Appendix F. Additional Information on Grades 7 – 8 Outcome Findings

Overall Relationship Between Average Weekly Lessons and Grade 7 – 8 Users' Outcomes on Math Inventory™

Table F1. Math Outcomes on Math Inventory™ by Weekly Lessons on DreamBox Math (covariates: beginning-of-year (BOY) Math Inventory™ score, grade, gender, race, ELL status, and IEP status)

Group Comparisons	Coefficient	Standard Error	t-value	p>|t|	Effect Size
<1.5 weekly lessons compared to 1.5–4.2 weekly lessons	10.69	4.56	2.35	.019*	0.07
<1.5 weekly lessons compared to >4.2 weekly lessons	30.07	7.96	3.78	<.001***	0.19
1.5–4.2 weekly lessons compared to >4.2 weekly lessons	19.38	8.58	2.26	.024*	0.14

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Overall Relationship Between Average Weekly Minutes and Grade 7 – 8 Users' Outcomes on Math Inventory™

Table F2. Math Outcomes on Math Inventory™ by Weekly Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, grade, gender, race, ESL status, and IEP status)

Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<14 weekly minutes compared to 14 – 38 weekly minutes	3.07	4.54	0.68	.499	0.02
<14 weekly minutes compared to >38 weekly minutes	19.31	6.70	2.88	.004**	0.13
14 – 38 weekly minutes compared to >38 weekly minutes	16.25	7.37	2.21	.028*	0.11

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Relationship Between Average Weekly Lessons and Grade 7 – 8 Users' Outcomes on Math Inventory™ by Grade

Table F3. Grade 7 Math Outcomes on Math Inventory™ by Weekly Lessons and Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, gender, race, and ELL and IEP status)

Kindergarten Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<1.5 weekly lessons compared to 1.5–4.2 weekly lessons	15.43	5.97	2.58	.01*	0.10
<1.5 weekly lessons compared to >4.2 weekly lessons	34.35	8.96	3.84	<.001***	0.22
1.5–4.2 weekly lessons compared to >4.2 weekly lessons	18.92	9.88	1.91	0.056+	0.14

Kindergarten Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<14 weekly minutes compared to 14 – 38 weekly minutes	3.20	6.02	0.53	.595	0.02
<14 weekly minutes compared to >38 weekly minutes	28.58	7.61	3.76	<.001***	0.19
14 – 38 weekly minutes compared to >38 weekly minutes	25.38	8.61	2.95	.003**	0.18

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table F4. Grade 8 Math Outcomes on Math Inventory™ by Weekly Lessons and Minutes on DreamBox Math (covariates: BOY Math Inventory™ score, gender, race, and ELL and IEP status)

Grade1 Group Comparisons	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect Size
<1.5 weekly lessons compared to 1.5–4.2 weekly lessons	3.76	7.02	0.54	.593	0.03
<1.5 weekly lessons compared to >4.2 weekly lessons	17.55	17.43	1.01	.314	0.13
1.5–4.2 weekly lessons compared to >4.2 weekly lessons	13.79	18.13	0.76	.447	0.11
<14 weekly minutes compared to 14 – 38 weekly minutes	2.91	6.86	0.42	.671	0.02
<14 weekly minutes compared to >38 weekly minutes	-12.07	14.36	-0.84	.401	-0.09
14 – 38 weekly minutes compared to >38 weekly minutes	-14.98	15.14	-0.99	.323	-0.11

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Relationship Between Average Weekly Lessons and Grade 7 – 8 Users' Outcomes on Math Inventory™ by Race, ELL Status, IEP Status

Table F5. Partial correlation coefficients and associated *p*-values for analyses by race, ELL status, and IEP status (covariates: grade, gender, BOY Math Inventory™ score)

	n	Partial Correlation Coefficient	*p*-value
Asian	52	0.16	.297
Black/African American	54	0.01	.961
Hispanic	225	0.11	.105
White	2,093	0.09	<.001***
ELL Status	47	0.37	.015*

IEP Status	363	0.05	.311

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Relationship Between Average Weekly Minutes and Grade 7 – 8 Users' Outcomes on Math Inventory™ by Race, FRL, and ELL Status

Table F6. Partial correlation coefficients and associated *p*-values for analyses by race, ELL status, and IEP status (covariates: grade, gender, BOY Math Inventory™ score)

	n	Partial Correlation Coefficient	*p*-value
Asian	52	0.12	.421
Black/African American	54	-0.11	.438
Hispanic	225	0.08	.211
White	2,093	0.05	.013
ELL Status	47	0.32	.041*
IEP Status	363	0.05	.364

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table F7. Comparison between matched samples of DreamBox Math users and non-users on end-of-year Math Inventory™ scores (covariates: grade, race, ELL status, BOY Math Inventory™ score, propensity score weights)

Outcome	Coefficient	Standard Error	*t*-value	*p*>|*t*|	Effect size
End-of-year Math Inventory™ Quantile® scores	17.66	3.41	5.18	<.001***	0.12

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$